DOCUMENT RESUME

ED 173 417 TM 009 461

AUTHOR Demaline, Randy E.; Quinn, D. William
TITLE Hints for Planning and Conducting a Survey and a Bibliography of Survey Methods. Aid #2 in Instructional Aids Series.
INSTITUTION Western Michigan Univ., Kalamazoo. School of Education.
PUB DATE Apr 79
NOTE 107p.
AVAILABLE FROM The Evaluation Center, College of Education, Western Michigan University, Kalamazoo, Michigan 49008 ($3.00)

EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Attitude Tests; *Data Analysis; Ethics; *Public Opinion; *Questionnaires; *Research Design; Research Methodology; Social Science Research; *Surveys; Test Construction
IDENTIFIERS *Mail Surveys

ABSTRACT

Methods of planning and administering mail surveys, developing questionnaires, and analyzing data are reviewed. Each review section is followed by an annotated list of selected readings. Topics discussed in planning a survey include decision making; survey designs; sampling plans; and ethical considerations. Developments of instruments and tupes of attitude measures are discussed in the section on survey instruments. Survey management is concerned with the mechanics of distributing and collecting the questionnaires, coding, and checking for errors. Data analysis focuses on nonresponse analysis, computer usage, and the choice of statistical methods. A 365-item bibliography and a subject index geared to the bibliography are appended. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Mary Louise Buenda

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED173417

HINTS FOR PLANNING AND CONDUCTING A SURVEY AND A BIBLIOGRAPHY OF SURVEY METHODS

April 1979

EVALUATION CENTER

COLLEGE OF EDUCATION
WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MICHIGAN 49008

TM 009 461

Aid #2
in
Instructional Aids Series

WESTERN MICHIGAN UNIVERSITY
KALAMAZOO, MICHIGAN 49008

INSTRUCTIONAL AIDS SERIES

The Instructional Aids Series is published by the Evaluation Center, Western Michigan University. It provides materials that are useful to those who teach evaluation theories and practices and to those who are engaged in independent studies of evaluation.

Additional copies may be obtained at a nominal fee to cover reproduction and mailing costs.

For information about other publications in this series, see the inside of the back cover of this document.

All correspondence in reference to the Instructional Aids Series should be addressed to:

Mary Anne Bunda
Editor, Instructional Aids Series
The Evaluation Center
Western Michigan University
Kalamazoo, Michigan 49008

EVALUATION CENTER

Daniel L. Stufflebeam
Director

Mary Anne Bunda

James R. Sanders

ABOUT THE CENTER

The Evaluation Center is committed to advancing the science and practice of evaluation. To serve its purpose the Center has developed components in research, development, instruction, dissemination, and service. Over the past ten years the staff of the Center has had significant experience in these areas working with school districts, state departments of education, foundations, and federal agencies.

The *Research Component* is designed to develop and test evaluation theory. Particular thrusts are testing the effects of school-based evaluation offices on school practice, and determining the special evaluation skill and knowledge profiles required by various groups in education. The *Development Component* is designed to work on ways and means of implementing evaluation theory. Included are techniques for meeting special evaluation objectives, new data collection instruments, and special procedures for communicating evaluative findings and relating them to practical problems. The *Instructional Component* is designed to provide in-service training in evaluation to various local, state, and national groups. While the Center is not a credit-granting agency, it collaborates with instructional units throughout Western Michigan University for the development and implementation of pre-service evaluation curricula. Emphasis in this work is at the doctors and masters levels. The Center serves as an instructional laboratory wherein students majoring and minoring in evaluation can gain practical evaluation experience. The Center is also committed to the development of novel and needed evaluation training materials and strategies. The *Dissemination Component* is designed to communicate research findings, new ideas, and materials to the education profession. *The Service Component* is designed to assist agencies in meeting their evaluation needs through two primary services. These are the design and implementation of evaluation studies and the development and installation of agency evaluation systems.

The staff of the Center has been involved in conceptualizing the CIPP Evaluation Model, researching the Item Sampling technique, developing the Advocate Team Technique, developing over 100 standardized tests, designing and conducting many in-service evaluation programs in evaluation. Other accomplishments include authoring the PDK book, EDUCATIONAL EVALUATION AND DECISION MAKING, and numerous journal articles and papers, designing and conducting many evaluation projects, and installing evaluation systems in schools, government agencies, and research agencies. Overall, the Center staff has directed externally funded projects in evaluation totaling more than four million dollars.

HINTS FOR PLANNING AND CONDUCTING A SURVEY AND A BIBLIOGRAPHY OF SURVEY METHODS

by

Randy E. Demaline

and

D. William Quinn

Western Michigan University

April 1979

Page

Preface i

PART I: PLANNING AND CONDUCTING A SURVEY

Introduction 1

I. PLANNING A SURVEY 2

Decision Considerations 2
- Appropriate Objectives
- Inappropriate Objectives
- Assumptions About the Respondent
- Factors Effecting the Cost of a Survey
- Selected Readings

Survey Designs. 7
- Cross-sectional
- Multi-wave
- Trend
- Longitudinal
- Combination
- Selected Readings

Sampling Plans. 14
- Nonprobability Sampling
- Probability Sampling
- Selected Readings

Ethical Considerations 20

- Selected Readings

II. SURVEY INSTRUMENTS 24

Development of Instruments 24
- Free-Response and Forced-Choice Questions
- Types of Forced-Choice Instruments
- Selected Readings

Attitude Measures 37
- Nominal
- Ordinal
- Interval
- Likert Scales
- Thurstone Scales
- Semantic Differential
- Guttman Scales
- Selected Readings

III. SURVEY MANAGEMENT 47

The Mechanics 47
Delivery of the Questionnaire
After the Delivery
Coding
Data Cleaning or Error Checks
Selected Readings

IV. DATA ANALYSIS 59

Nonresponse Analysis 59
Selected Readings

Using the Computer 60
Selected Readings

Choosing Statistics 62
Nominal
Ordinal
Interval
Measures of Association Between Two Variables
Selected Readings

PART II: BIBLIOGRAPHY OF SURVEY METHODOLOGY

I. BIBLIOGRAPHY 70

II. INDEX TO BIBLIOGRAPHY 94

Preface

From time to time the Evaluation Center receives request for information about questionnaires--their construction, format, analysis, etc. It was because of these requests that we started work on a bibliography about questionnaire construction. However, it soon became apparent that one cannot isolate the construction of the questionnaire from some other areas of survey methodology. In addition to expanding the scope of the bibliography, it was also decided to add a narrative, briefly sketching the phases of survey methodology plus sensitizing the reader to some of the foreseeable decisions that might have to be made. The goal of this completed document is for the reader to be able to obtain summaries of some of the important topics in survey methodology and also be able to more efficiently locate literature for more indepth information.

The document is designed to be of a supplemental nature. It surely has not been written to replace methodological coursework or the several good textbooks or handbooks written on the topic. The narrative is written for the reader who has had little or no introduction to survey methodology. The more advanced practitioner will find the bibliography of most use.

To limit the scope of the narrative, it was decided to deal mainly with questionnaires and, even more specifically, questionnaires in mail surveys. Field interviews or phone interviews or forms handed out in group settings, a large and important area of survey methodology, is only lightly touched upon. Therefore, most of the summaries and the majority of the references are concerned with conducting a mail survey, and the topic of interviewing techniques is not covered.

The reader will find that the document is divided into two main parts: Part 1, a summary of survey methodology as we have limited it and Part II, a bibliography. In part I, major topics in survey methodology are briefly discussed. Major sections conclude with annotated lists of selected readings, listing those references that either provide a more thorough treatment of the topic or provide excellent examples. When larger treatises are referenced, the reader is directed to relevant chapters or sections and provided with our judgment of the level of difficulty of the material.

Part II of this document is comprised of the bibliography and a topic index. Entries in the bibliography are sequentially numbered and these numbers are indexed to relevant topic areas. Again, we make no claims for the bibliography's comprehensiveness. The number of entries seems few when compared to, for example, the U.S. Department of Commerce's (1974) nearly 2,500 entries. We have, however, tried to include what could be considered the classics along with some of the more recent materials to appear on survey methodology. There are probably many alternative readings that could have been suggested, the exclusion of a reference to any particular work should not be interpreted as a comment on its comparative merit.

The wide range of references should not imply that we suggest the reader should have expertise in all areas of survey methodology before embarking on a study. To infer that survey methodology should be limited only to trained statisticians or sampling theorists would be an unfortunate conclusion.

Finally, in the tradition of prefaces, we wish to give grateful acknowledgement to the Evaluation Center at Western Michigan University

for providing financial and moral support. Particular thanks are due Associate Director Mary Anne Bunda whose constant encouragement guided us in our task.

D. William Quinn

Randy E. Demanline

PART I

Planning and Conducting a Survey

The narrative presented in this part of the document is divided into four major sections: Planning a Survey, Survey Instruments, Survey Management and Data Analysis. Each of these sections contain short descriptions of various decisions the potential surveyor will make. We do not expect a reader to procede through the narrative and then know how to conduct a survey. Rather we suspect that the reader has a question in one of the four major areas, e.g., 1) Is a survey what is really needed for this investigation and how should it be designed?, 2) How do you construct a survey instrument?, 3) What are some of the management problems in survey research?, or 4) How do you analyze this particular kind of data?

Each of the four major sections is divided into subsections so that the readers may more quickly find the answer to their particular concern. These subsections are neither mutually exclusive, nor exhaustive, but rather are the concept areas in which we most often get questions. We have tried in the narrative to be consistent in the level of discourse. It is easy when handling a request in person to search out a level of communication. Unfortunately, the level of communication in a print medium cannot be negotiated, and the reader may find some slippage. Moreover, we have not answered some specific questions in the narrative. For this reason each subsection is concluded with a list of alternative sources for information.

The reader should also be aware that the discussion in each section has been limited by the decision to keep this document relevant to mail questionnaire surveys only. While some of the discussion, e.g. Analysis, may also be applicable to more general cases of inquiry, we have not tried to be comprehensive in those sections. Nor have we contrasted the objectives of mail surveys with those of other inquiry methodologies. The

discussion subsequent to "Should we do a mail survey?" will be helpful for those people who decide "yes." But it will not provide help to those individuals who decide that a mail survey is not appropriate for their inquiry. We, also, have used the word "survey" in this restrictive sense without always using the appropriate qualifiers. The sophisticated reader will see much simplification in the narrative. However, we hope that our simple view of the process is not perceived as an "over simplification."

I. PLANNING A SURVEY

Decision Considerations

In any study, it is important that there be a match between the types of questions being asked, the types of data being collected, and the methods for collecting those data. If the investigation asks the question "What are the opinions of a group of people on certain issues?", then the investigator is interested in gathering descriptive data. However, if the question is what caused people to arrive at their current opinions, then more than descriptive data may be required to identify cause and effect. In general, deciding whether a survey is a viable approach for your study involves addressing six kinds of concerns:

(1) the kind of objectives and the precision needed to address them,

(2) the type of information required,

(3) the type of respondent needed to supply the information,

(4) the accessibility of these respondents,

(6) the resources (human and dollar) available, and

(7) the ethical consideration.

The rest of this section will briefly elaborate on these concerns.

Appropriate Objectives. A survey can provide appropriate data for the following types of objectives:

(1) The description of a group. This description may be done by considering variables singularly (e.g., describing the average score or spread of the scores). Or the description may be accomplished through comparisons and contrast by examing the relationships between the variables (e.g., opinions and purchasing preferences), or by comparing various young respondents, large organizations versus small organizations, actual versus desired responses).

(2) The exploration of possible cause and effect relationships. With some survey methods it is possible to obtain limited information about the naturally occurring causal relationship between variables. This is usually done by observing a correlational relationship between variables, developing theory to explain this correlational relationship as indicating a causal relationship and "confirming" theory through such correlational studies as path analysis or cross-lagged panel correlation while eliminating plausible vital explanations for the correlational relationship. (See Glock, 1967, Blalock, 1964, or Rosenberg, 1968, for an explanation of surveys for causal analysis. See Campbell and Stanley, 1966, or Cook and Campbell, 1976, for a discussion of the limitations of such correlational designs.)

Inappropriate Objectives. A survey cannot provide appropriate data for the following types of objectives:

(1) The description of a complex social process. Surveys indicate the

disposition of respondents at one or more points in time, but they are not useful for indicating how or why the respondents arrived at their standing. Various case study and observational techniques are more effective for describing the complex interaction of people and organizations than are survey techniques.

(2) The identification of cause and effect relationships. Although some survey designs may be useful for identifying possible causal relationships, they are generally less suited for testing these relationships. Experimental and quasi-experimental designs are better suited to investigate causal relationships.

Assumptions About the Respondent. In addition to insuring that your questions can be answered by survey data, it is important to consider assumptions that you will be making about the respondents. Some of these are that the respondent:

(1) is knowledgeable enough to provide the information you want.

(2) is willing to provide this information.

(3) understands the survey questions as you intend.

(4) answers the questions in the form you want.

(5) will not lie or otherwise subvert the survey process.

(6) will not be harmed by the survey process.

(7) will not consider the survey an invasion of privacy.

You will also be faced with identifying the group of respondents who will be able to supply the information you need to answer your hypotheses plus the logistics problems of obtaining access to them. If the assumptions listed above are not tenable, the quality of the information gathered is suspect.

Factors Effecting the Cost of a Survey. When choosing whether to do a survey, be sure to consider costs. Anyone who has done a survey will verify that surveys are not to be found in the bargain basement of investigatory techniques.

Listed below are possible expense categories to consider:

Personnel

Project Director
Study Supervisor
Research Specialist
Secretary
Consultant

Services

Travel
Coding
Preparation for machine analysis
Data processing
Keypunching
Typing

Supplies and Materials

Paper
Envelopes
Duplication
Printing
Data cards or tapes
General office supplies
Communication
  Postage
  Telephone
Rentals

Other

Institutional overhead
Employee benefits

Doing a survey, then, depends on whether you have enough money, whether the respondents are likely to cooperate if you behave in an ethical manner and whether your inquiry questions are answerable by this methodology. If you are unsure of the applicability of mail survey methodology to your inquiry, more in depth reading is necessary.

Selected Readings.

Babbie, E.R. Survey research methods. Belmont, California: Wadsworth, 1973.

This text covers some of the basic skills and methods involved in survey research. Also topics that are not covered in many other basic tests, such as ethics of survey research, are included in this text. Chapters 1 through

4 deal particularly with the scientific context of survey research and the types of survey designs.

Backstrom, C.H., & Hursh, G.D. Survey research. Evanston, Illinois: Northwestern University Press, 1963.

Chapter 1 contains a basic discussion of what a survey using field interviews can and cannot do. Also, a discussion of survey costs is included.

Hyman, H. Survey design and analysis. New York: Free Press, 1955.

The topics of descriptive surveys and explanatory (causal) surveys are well developed in this text. A generous number of case studies of surveys are provided to illustrate general principles discussed.

Oppenheim, A.N. Questionnaire design and attitude measurement. New York: Basic Books, 1966.

This text covers the several phases of questionnaire research. The pros and cons of several approaches to survey research are discussed. Chapter 1 is primarily concerned with survey design.

Survey Designs

As investigator's questions may vary, the particular designs involved in surveys are also varied. Five basic survey designs are discussed in this section: cross-sectional, multi-wave, trend, longitudinal, and the combination of a survey with other research methods. These designs vary in the number of measurement periods involved, the types of variables considered, and the samples measured. Each design is useful for providing different information according to the objectives involved. This list of designs is by no means exhaustive, but demonstrates the variety of designs that may be used. If these designs do not match your objectives, consult a text book.

Cross-sectional. A cross-sectional survey involves measuring a sample from a population at one point in time. Figure 1 shows the basic idea of this design. As in most surveys, the measures may involve factual, attitudinal, or judgmental questions as well as demographic information. The most common use for a cross-sectional survey is to an inferences about the population from which the sample was drawn. This survey design is well suited to this inferential purpose. However, some tenative causal inferences may be made from cross-sectional surveys. These exploratory causal investigations may be made when questions are included in the survey that can be related to different time periods in the subject's life. When a substantial relationship is observed between variables from different time periods, this may indicate that differences in the earlier variable caused changes in the later variable. For example, if on a survey of employee success and previous training it was

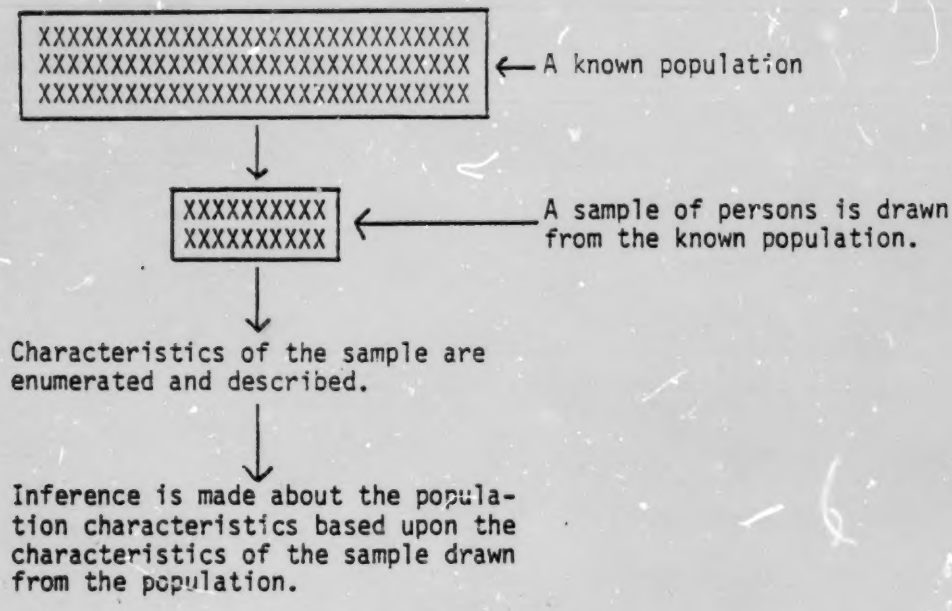
XXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXX
A known population
XXXXXXXXXX
XXXXXXXXXX
A sample of persons is drawn from the known population.
Characteristics of the sample are enumerated and described.
Inference is made about the population characteristics based upon the characteristics of the sample drawn from the population.

Figure 1

The Cross-Sectional Design[a]

[a]Based on Black and Champion (1976).

found that participation in a particular training program was highly related with later job success, it may be inferred that learning gained in the earlier training program produced later job success. But since other factors, such as general ability, may account for both being admitted to the training program and job success, it is not reasonable to conclude on the basis of a cross-sectional survey along that a causal link exists between these variables. This type of survey can only indicate possible causal relationships and other methods must be used to validly test these relationships. Conclusions about cause on the basis of a single cross-sectional design are likely to be based on a weak data source.

Multi-wave. A multi-wave survey involves measuring one sample of subjects at two points in time on related but different measures. This design may be used to identify otherwise hard to locate sub-groups of a population for more intensive study. This would involve a screening survey followed by a more extensive survey for subjects belonging to the sub-group of interest. See Figure 2 for a pictoral representation.

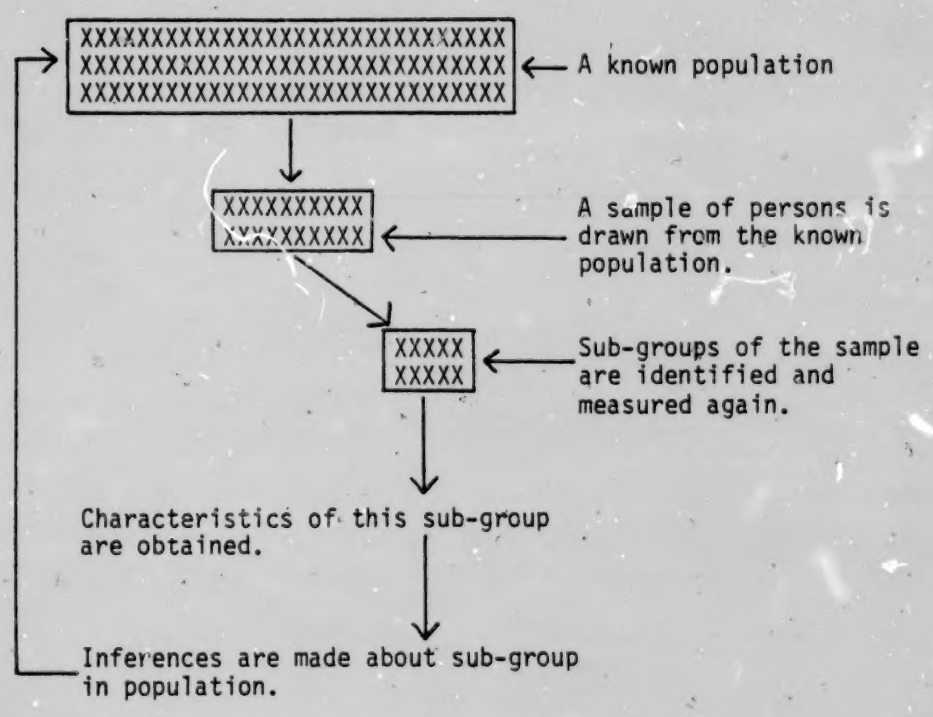
A known population
A sample of persons is drawn from the known population.
Sub-groups of the sample are identified and measured again.
Characteristics of this sub-group are obtained.
Inferences are made about sub-group in population.

Figure 2

The Multi-Wave Design

Multi-wave surveying may also be useful when information from the previous survey is used as part of the questions on later surveys.

Trend. A trend survey involves measuring different samples from the same population on the same variables at two or more points in time. Figure 3 demonstrates a trend survey.

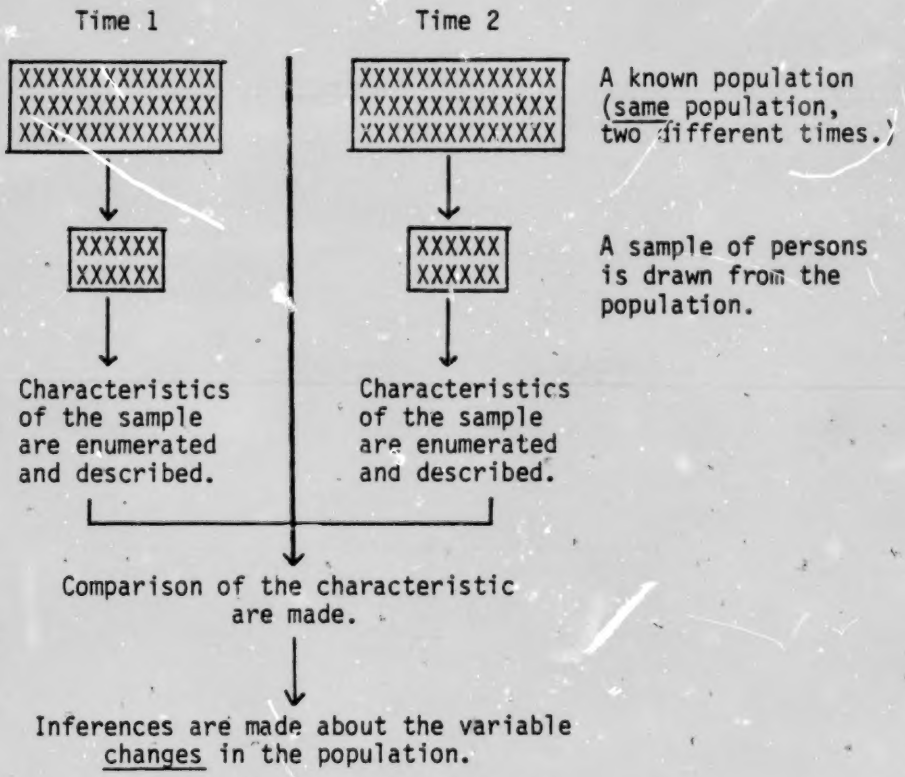
Time 1
Time 2
A known population (same population, two different times.)
A sample of persons is drawn from the population.
Characteristics of the sample are enumerated and described.
Characteristics of the sample are enumerated and described.
Comparison of the characteristic are made.
Inferences are made about the variable changes in the population.

Figure 3

The Trend Design

Trend studies can describe changes that occur in a population over time. In this sense they provide a fuller, more dynamic description of the population than cross-sectional surveys. They also can provide somewhat stronger evidence about cause since time with respect to one variable change against another can be more directly observed. However, unconsider

or unmeasured factors can still provide plausible rival hypotheses for changes. Changes in the sample being studied, history and other factors may cause one variable measured at an earlier time to appear to cause the change in a second variable measured at a later time. For trend surveys, as with other survey strategies, an indication of possible cause and effect relationships can be obtained, but not a strong demonstration of such a relationship.

As Figure 3 shows the sample is drawn from the population independently at each point in time. Consequently, the design is equivalent in operation to two cross-sectional designs separated by a period of time. However, the intent of the data collection is quite different in that the question of interest is change in the population and not description of the population at any particular point in time.

Longitudinal. A longitudinal survey involves measuring one sample on one set of variables at two or more points in time. Figure 4 demonstrates this. This survey strategy provides the best opportunity of the several strategies discussed for demonstrating causal relationships between variables. (See Campbell, 1963; Pelz and Andrews, 1964, for a description of an analysis technique which utilizes panel data for causal inferences.) These causal analyses of longitudinal surveys are limited to describing causal relationships that acutally occur and not those effects that could possibly occur if other "causes" were actively introduced by the researcher. This design also may identify causal relationships without indicating whether the causes are amenable to manipulation. Therefore, the researcher may not be able to manipulate a particular effect.

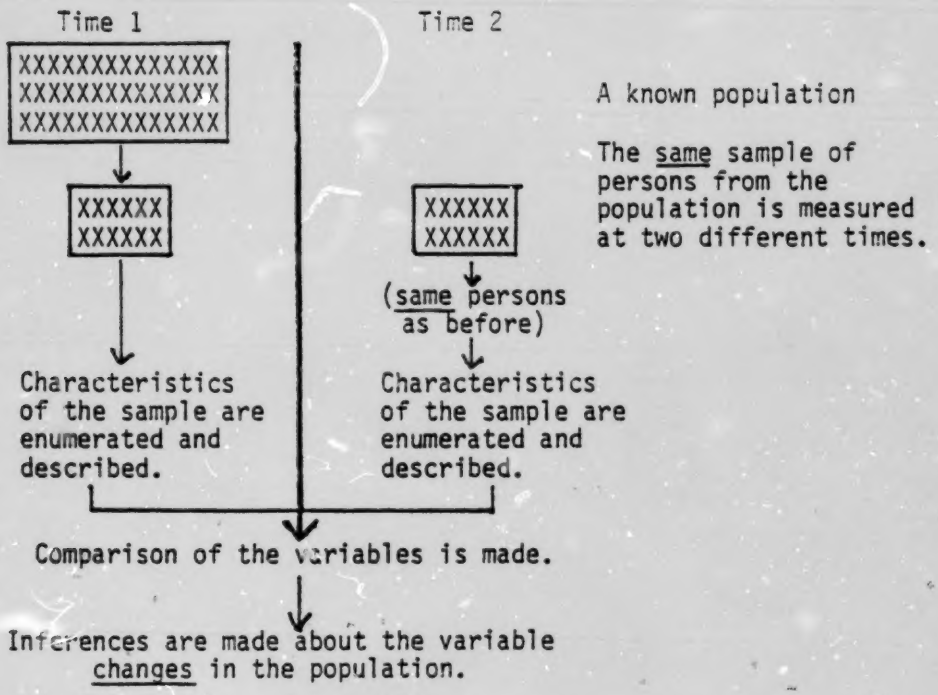
Time 1
Time 2
XXXXXXXXXXXXXX
XXXXXXXXXXXXXX
XXXXXXXXXXXXXX
XXXXXX
XXXXXX
XXXXXX
XXXXXX
A known population
The same sample of persons from the population is measured at two different times.
(same persons as before)
Characteristics of the sample are enumerated and described.
Characteristics of the sample are enumerated and described.
Comparison of the variables is made.
Inferences are made about the variable changes in the population.

Figure 4

The Longitudinal Design

Note that the longitudinal design involves the measurement of a single sample of persons at two different times. Some researchers refer to this sample as a "panel."

Combination. By combining survey strategies with other research methodologies, such as case studies or experimental studies, it is possible to produce hybrid designs which are better suited to answer a research question than any of the methodologies are in isolation. This technique is often referred to as trianglation. Webb, et al., (1966) in the book Unobtrusive Measures: Nonreactive Research in the Social Sciences have suggested several novel measurement techniques that may be used with survey methods to their mutual strengthening. The methods suggested included measures involving physical traces, archives, and observations.

Their proposition is that multiple techniques allow for richness of information and strength of conclusion.

Selected Readings.

Babbie, E.R. Survey research methods. Belmont, California: Wadsworth Publishing Company, 1973.

Once again, this textbook offers a comprehensive introduction to the areas of survey methodology with Chapters 4 and 6 explaining types of study designs and providing examples.

Black, J.A., & Champion, D.J. Methods and issues in social research. New York: John Wiley & Sons, 1976.

This is a general textbook on research methods in the social sciences. It gives a good explanation of research designs and the advantages and disadvantages of each.

Glock, C.Y. (Ed.) Survey research in the social sciences. New York: Russell Sage Foundation, 1967.

The application of survey methods in the various fields of sociology, political science, psychology, economics, anthropology, education, social work and public health are discussed in this volume. Pages 3 through 58 contain a particularly useful description of survey design strategies.

Sampling Plans

The major ingredients of a survey are a questionnaire, a data analysis plan, and a sampling plan. The discussion which follows draws heavily on "Sampling Options for School Research and Evaluation," a paper by James Sanders presented at the Washington Research and Evaluation Conference, May, 1974. Sampling involves drawing inferences about a population of interest based upon measuring only a part of the population. The population is defined by the researchers so that the questions under investigation can be answered. A population of a surveyor does not mean the same thing as the "population of a city." For social science or educational surveys, the populations are usually made up of people or organizations.

The particular elements that make up a population are called the sampling units. Care must be taken that the population and sampling units actually measured are congruent with those needed to answer the questions of interest. The unit of sampling has great importance since it also becomes the unit of analysis. If the survey questions concern individuals, then it would be inappropriate to sample and analyze on a group level. Additionally, once the sampling is completed the analysis is virtually determined by the plan which has been implemented. If the survey question involves description of the organizational climate of medium size stores in a given city, it is important that only stores of that size be included in the sample. It is also important that the appropriate employees in these stores respond to the survey and not just some sub-group such as managers or new employees (appropriate is indicated by the definition of employee used for the study).

To help identify and select sampling units, a sampling frame is identified or constructed. The sampling frame is a map, directory, lists, or procedured for identifying sampling units for possible inclusion in the survey. If you were interested in surveying all legally practicing physicians in a state, then the sampling frame might be a list of all licensed physicians from the state regulatory agency. Other individuals who practice medicine would not be part of the population, since the legality of practice is part of the definition of the population.

For samples to be useful, the units sampled must reflect the population they are part of within a given degree of accuracy. Different models have been used for drawing a sample from a population. These models are nonprobability sampling (including haphazard, judgment, and quota sampling), and probability sampling. Each of the three types of sampling discussed in the next section represent ways of selecting units in a non-scientific manner. These sampling procedures have grave consequences for the analysis and interpretation of the resultant data. Probability samples, on the other hand, are easier to analyze and interpret since the degree to which they represent the population is known.

Nonprobability Sampling. *Haphazard* sampling involves selecting a sample based on convenience without any concern for how well it represents the population. When someone selects all the people in a convenient apartment house for a survey of attitudes of people in a city, then this is a haphazard sample or sometimes called a "grab" sample. This type of sample is generally less costly, less difficult, and less time consuming to conduct than other models, but the results are not very representative of the population.

Judgment sampling involves selecting a sample by some, hopefully, knowledgeable person to represent the population. If a city planner chooses four communities which she thinks represents all of the people in the city, then she has selected a judgment sample. Judgment sampling is, in general, moderately costly, difficult, and time consuming. The results are likewise moderately representative of the population.

Quota sampling involves drawing a sample based upon specified quotas. The quotas may be set to represent some aspect of the population, such as age or race. Subjects are then categorized into one or more of the quota categories and surveyed if the category (ies) is not already filled. If through Census returns it has been determined that citizens of a city were divided between blacks and whites in a one to ten ratio, and also that each race were evenly divided between males and females a surveyor may decide not to use scientific designs because blacks maybe only a small part of the sample. If the researcher wants to be sure that the races and sexes are measured he may set quotas of 50 black males, 50 white males, 50 black females, and 50 white females. The first 200 people that fill these categories would be used to represent the city's citizens. Like judgment sampling, quota sampling generally is moderately costly, difficult, and time consuming; the results generally are not representative of the population. In the example above the quotas were set without regard to the population characteristics, consequently the sample does not represent the population, even though the equal size of the sub-groups does make analysis easier.

Probability sampling involves assigning every unit in the sampling frame a non-zero probability of being selected and then selecting a sample

based on these probabilities. There are several variations on this model and they will be discussed shortly. Probability sampling generally is the most costly, difficult, and time consuming model; however, the results represent the population to a known degree, i.e., the amount of possible error can be estimated. Hence, the surveyor can make statements of confidence concerning the accuracy of the information collected. With the non-probability sampling models, no convenient way exists for judging the degree to which the sample represents the population.

Probability sampling is clearly the best sampling model for survey investigations. Because many excellent textbooks exists on this subject, and because of the complexity of the subject, only some basic variations on probability sampling are defined below. The reference list at the end of this discussion provides sources for further treatment of survey sampling. The four variations on probability sampling discussed here are simple random sampling, cluster sampling, stratified sampling, and multi-stage sampling. These are by no means mutually exclusive in that you will find many large studies that perform a combination of these in their sampling plans.

Simple random sampling means each unit in the frame has an equal probability of being selected. If all citizens of a city are listed somewhere (a rare event) and 200 of them are selected from the list randomly (i.e., each person has an equal chance of being selected), then this is simple random sample. The procedure usually involves assigning numbers to the citizens and then using a random number table to select from them. (Many textbooks have reproduced parts of the random number tables published by the Rand Corporation, 1955.)

Another technique that is often used and usually produces equivalent results is *systematic* sampling. If, for example, a list of names were produced and 20,000 names were on the list, systematically sampling every 100th person on the list, given a random starting point in the first 100 people, would probably be equivalent to a truly random sample of 200. The major caution to be aware of with this systematic sampling is that there is no cyclical order to the list. For example, if for some reason, all odd numbered persons were males and all even numbered persons were females, selecting every 100th person would result in surveying only males or females. In this case, systematic sampling would not produce equivalent results to simple random sampling.

Cluster sampling involves randomly selecting some naturally occurring cluster of sampling units and then surveying all of the units in that cluster. Randomly selecting city blocks and then interviewing all of the residents of these blocks would be cluster sampling. Cluster sampling is usually done because of the greater ease of surveying all units in a cluster rather than random sampling on an individual basis. Particularly if the units in the sample are geographically spread out and interviews or other site data collection methods are planned as the data collection plan.

Stratified sampling is used so that different sub-groupings in the population are taken into account in the sampling. In a national survey of opinions where regional differences are likely to occur, the sample would stratify the country by geographic region and within each strata a random sample would be selected. This technique is often done to ensure that groups are present in sufficient size to provide inter-group comparisons. Of course, characteristics of the population must be well known in order to use this procedure. In other words the relative

proportion of the number of units in each strata must be known. Also the variable used for stratefication, e.g., geographic region or race, should be relevant to the questions the survey is to investigate.

Multi-stage sampling involves randomly selecting some intermediate sampling units, and then drawing the final sample randomly from within these units. If city blocks were randomly sampled and then a sample of residents in each block were randomly sampled, this would be multi-stage sampling. Multi-stage sampling has the advantage of not needing to list each of the final units in the sampling frame before making the selection.

Selected Readings.

Kish, L. Survey sampling. New York: Wiley, 1965.

This is one of the classic books on sampling, but is not recommended to anyone who does not have a good mathematical background.

Sudman, S. Applied sampling. New York: Academic Press, 1976.

This book was written for the much less mathematically sophisticated than the above Kish reference. It also provides good examples and practical considerations for alternative sampling procedures.

Ethical Considerations

The public recently has become more aware of the effects of studies on their participants and the rights of these subjects. Often discussions in this area deal with two issues: the protection of human subjects and the participants' rights to privacy. The first of these, protection of human subjects, deals with avoiding causing the respondent physical or mental stress. The second deals primarily with the respondent's right to refuse to be involved in a study and the unethical collection of data about a respondent without his/her knowledge or consent.

Since the enactment of federal laws requiring the review of federally sponsored research and development projects, almost all universities and many other public and private institutions have established policies and procedures for the internal review of projects involving human subjects. These procedures are designed to discover any possible physical, psychological, or social harm to the participant as a consequence of participation in the study. The ultimate responsibility, however, lies with the investigator.

One method commonly employed to safeguard a subject's rights and welfare is to incorporate informed consent as a prerequisite to a person's participation in a project. The meaning of informed consent is presented in the following definition given in the Protection of Human Subjects Rules and Regulations (Department of Health, Education, and Welfare, 1975).

> "Informed consent" means the knowing consent of an individual or his legally authorized representative so situated as to be able to exercise free power of choice without undue inducement or any element of force, fraud, deceit, duress, or other form of constraint or coercion. The basic elements of information necessary to such consent include:

> (1) a fair explanation of the procedures to be followed and their purposes, including identification of any procedures which are experimental;
>
> (2) a description of any attendant discomforts and risks reasonable to be expected;
>
> (3) a description of any benefits reasonably to be expected;
>
> (4) a disclosure of any appropriate alternative procedures that might be advantageous for the subject;
>
> (5) an offer to answer any inquiries concerning the procedures; and
>
> (6) an instruction that the person is free to withdraw his consent and to discontinue participation in the project or activity at any time without prejudice to the subject. (pp. 11854-11855).

Some of the particularly relevant aspects of informed consent for survey research are: (1) the voluntary participation of subjects without deceit or other coercion; (2) the disclosure of the purpose of the survey and its sponsor, and (3) the description of possible discomfort or risks resulting from answering the questions. This third point is particularly relevant when subjects are asked about opinions or activities that could be socially embarrassing or illegal (e.g., sex habits, use of drugs).

The use of confidentiality or anonymity in connection with survey responses is also a common practice for protecting the rights of participants. Confidential data collection procedures are those in which the investigator is able to identify which responses are from which respondents but promises not to reveal this information to others. Anonymity, on the other hand, is the arrangement whereby not even the investigator knows which person submitted a given response. Various

procedural and statistical techniques can be used to help assure that information given will remain confidential or anonymous (Boruch and Cecil, 1978). Because of the complexity and importance of this issue, you are advised to consult the publications in the following list of readings for a full understanding of how protection of human subjects relate to your survey.

Selected Readings.

American Association for Public Opinion Research. Code of professional ethics and practices. Public Opinion Quarterly, 1960, 24, 529-530.

This article would be useful for anyone considering a questionnaire as a research strategy. It will alert you to any possibilities of possible unintentional but unethical practices.

Boruch, R.F., & Cecil, J.S. Methods for assuring confidentiality of social research data. Book in preparation, 1978.

Procedural, statistical, and legal methods for protecting confidentiality both in the original research and during secondary analysis of existing data banks are discussed. This will be a very worthwhile publication.

Department of Health, Education and Welfare. Protection of human subjects (Technical amendment). Federal Register, 1975, 40(50), part 2, 11854-11855.

Provides the HEW rules and regulations regarding human subjects review. It is a good guide for constituting a review committee and basic definitions and policies for the review.

Michael, J.A. & Weinberger, J.A. Federal restrictions on educational research: Protection for research participants. Educational Researcher, 1977, 6(1), 3-7.

Michael, J.A., & Weinberger, J.A. Federal restrictions on educational research: Privacy Protection Study Commission hearing. Educational Researcher, 1977, 6(4), 15-18.

Weinberger, J.A., & Michael, J.A. Federal restrictions on educational research: A status report on the Privacy Act. Educational Researcher, 1977, 6(2), 5-8.

Weinberger, J.A., & Michael, J.A. Federal restrictions on educational research: A status report on the Buckley Amendment and Freedom of Information Act. Educational Researcher, 1976, 5(11), 3-8.

This series of four articles deals primarily but not exclusively with the Federal regulations involving protection of citizen rights in regard to research activities.

II. SURVEY INSTRUMENTS

More than other areas in survey methodology, writing the instrument is an art and science learned through experience rather than by precept. As an art many rules of thumb have been suggested, rules that have been only partially confirmed by research or organized by any theory of questionnaire writing. This section will aid you in writing questionnaires by outlining the general steps involved in questionnaire development, and referencing those sources where a large number of rules have been given or where good examples of questionnaires can be found. The science of instrument development, or knowing when the instrument produced data in which you can be reasonably confident, is also reflected in this section.

Development of Instruments

The five major steps in instrument development are:

1. Plan the questionnaire to respond to the issue and topics which are the central concern of the survey. This planning is done through the use of a questionnaire blueprint or outline. The blueprint should specify the topics, issues, hypotheses or concerns to be covered, the types and an appropriate number of questions for each subarea, as well as constraints on the instrument (e.g., limited time or number of pages) and special characteristics of the respondents (e.g., reading level). In other words the blueprint specifies what should be as well as what should not be part of the instrument.

2. Write the questions specified in the blueprint. As mentioned before, this is an art, as well as a science. Many rules of thumb

exist for guiding and checking your work at this stage. Payne (1973) has given 100 considerations to aid in writing questions and Allen (1957) has listed some informal criteria for writing attitude questions. Both of above sources are highly recommended. Below are some questions you should be asking yourself during the construction of the questions for the survey and place them together to form the questionnaire:

Some Guidelines for Questions in a Questionnaire:

1. The need for the question is clearly evident in the questionnaire blueprint.
2. The question is the most effecient way of gathering the information from the respondents, i.e., the information does not already exist in some other place.
3. The question contains only one idea and the respondent will be able to pick out the key word(s).
4. The question provides information which can be interpreted directly or it is followed by a probing sub-question which can be used to help interpret the responds.
5. The question concerns information which the respondents can supply, i.e., they know the answers or have opinions.
6. The question concerns a matter which will not irritate or embarrass the responderts.
7. The question is clear and direct so that respondents do not suspect ulterior motives on the part of the surveyor.
8. The language of the question is fair for all respondents, i.e., particular connotative meaning does not color the responses of any part of the population.
9. The question contains language which is likely to be familiar and clear to all respondents, i.e., to the extent possible all technical terminology has been removed.

10. The question is not so broad that the respondents can answer at various levels of specificity.

11. The question is independent of those which are before and after it in that previous questions are not leading. That is, the order of the questions on the questionnaire does not lead the respondents to answer in a particular way. The concept of independence does not apply to probe questions, but leading does.

12. The order of the questions on the instrument are in some logical or arbitary order, i.e., cronological or alphabetical.

13. The order of the questions on the instrument maintains interest and motivation so that the respondents will complete the questionnaire.

Some of the guidelines may appear contradictory. That is, in trying to keep the order interesting and motivating, a surveyor may put the items in an order that leads the respondent to a particular set of conclusions. However, the consequences of not thinking about the trade-off are more devistating. Which is a more critical error--possibly boring some respondents leading to low response rates or maintaining high interest while influencing the way in which people react? A surveyor must compromise to get the best possible information from the most people.

When you write questions, it is usually wise to write more than you need so that you can afford to discard items or formats that don't work in the pilot testing for one reason or another. This is particularly true when attitudes are to be measured as part of the questionnaire. In these cases, you should write perhaps twice as many questions as you will finally need. As can be seen by the last several guidelines, the way in which the questions are presented to the respondent is an important concern. Hence, care should be taken in the format of the questionnaire.

3. Format the questionnaire with these three criteria in mind: attractiveness, clarity for use, data coding. The questionnaire should be appealing to look at (at least not crowded or smudged) and reproduction on the best paper and with the most professional printing methods consistent with the budget for the survey. Brief directions for answering the questions and what to do with the form when answered should be clearly stated at the start of the questionnaire. Grouping similar items or items with the same response options ease the respondent's task. Number the questions and pages so that respondents don't inadvertently skip items. Threatening as well as dull questions at the beginning of the instrument may discourage people from answering the form. If placed at the end of the instrument, the respondents are likely to continue since they have already invested in filling out the questionnaires.

If you will be keypunching data for analysis, consult with the staff at the computer center you will be using for suggestions on how to make the form easy to code. Some general suggestions might include grouping together all questions with the same response format and placing answer spaces on the right side of each page in columns so the coder won't have to hunt for the data. However, the response form will complicate this issue. The main point here is that you should consider the coding of the information before you send out the forms.

4. Pilot testing of questionnaires is an essential part of instrument development. By the time a questionnaire is written, considerable thought and attention has already been given to the quality of the instrument. However, it is not unusual to still have ambiguous, confusing, or biasing questions or directions in the questionnaire. One way to

eliminate these unwanted elements is to pilot test the instrument. Pilot testing is usually done with a sample of subjects as similar to the true population of study as possible. These people are given the instrument and asked to respond as if they were actually being studied. These individuals provide a check not only on the items on the questionnaire but also the directions. Some may also be asked to annotate the questionnaire indicating words or questions they found unclear, offensive, or otherwise a problem. These people may also respond to open ended questions and their responses could then be used to produce categories to develop closed-ended questions for the full survey. Based upon the information from the pilot test, the instrument can be revised. For instance, items which have illicted irrelevant responses or items which have consistently been left blank in the pilot test need the attention of the author(s) of the instrument. It is much better to find a "red flag" in a small pilot than in the full survey.

5. Establishing a plan to check on the reliability and validity of the instrument. Reliability and validity refer to two simple ways of judging the quality of the data produced by the instrument. Reliability refers to the consistency of the information, while validity refers to the level of certainty that the instrument measures what you as the constructor thinks it does. Unfortunately, reliability and validity aren't qualities that can be wished onto a data set provided by an instrument. Surely careful attention to wording to obtain consistent respondent interpretation of questions and the reviewing of similar studies that have used similar techniques will help. There are no substitutions, however, for collecting data in your own study to judge the data's quality as you have used the instrument. Remember also,

that reliability and validity are not inherent to the instrument. They depend also on the use of the instrument. Therefore, even if you are using an instrument that has supporting reliability and validity information in some situations, you still must ask yourself about the effects of your target population, of the time that the instrument is being used, and of the conditions under which the population is being studied, and the quality of the data.

Reliability and validity can never be stressed too much. Unreliable instruments produce data so full of inconsistency and error that nothing definite can be stated. Data from a reliable but invalid instrument will provide data that is useless to your study. If you don't plan to collect these quality indicators concerning your data, you won't have it at the end of the study to temper the interpretation of the results.

Free-Response and Forced-Choice Questions

Questionnaire items are commonly categorized in two categories: free response (open-ended) or fixed response (forced-choice). A free-response question requires the respondents to answer in their own words. The interpretation, and response length and content are left to the respondents, since they are "free" to supply whatever they wish. Although, the length of the line(s) provided often give a hint to the respondent as to how long a response is anticipated. Examples of free-response items are:

What is the most serious problem in the high school(s) in your community? _____

What action should be taken by the school to improve high school education in your community? _____

_____

_____

A forced-choice question, on the other hand, provides respondents with two or more options to choose among in answering the question. You determine as the instrument author what aspects of the question the respondent will address when answering the question and limit the response choices. An example of a forced-choice item is:

> What is the most serious problem in the high school(s) in your community? (Circle one letter only)
>
> a. classroom discipline
> b. vandalism and violence
> c. substance abuse (drug, alcohol)
> d. academic achievement
> e. career preparation

When deciding which format to use, you must be aware of the advantages and disadvantages to both types of response formats. Some of these are listed below:

Free-Response Questions

Advantages

1. It can be easily formulated without knowing the full range of answers that may be given.
2. It can accommodate questions for which a wide range of different answers will be given.
3. It does not condition or bias the answer as much as the forced-choice question.

Disadvantages

1. It requires the respondent to write a lot. Communication skills may influence the answer, in addition the other respondent characteristics.
2. The respondent may address different facets of the question in which you may not be interested or may not give complete information in answering the question.
3. Fewer questions can be asked in a questionnaire because answers may be lengthy and time consuming to give.

4. Responses are difficult to analyze. The investigator must devise a coding scheme and then categorize responses based on this scheme. The diversity and richness of responses are usually reduced by this process, and it is time consuming.

Forced-Choice Questions

Advantages

1. It is easier for respondent to answer.
2. Focuses respondent's answers on issues and data of importance to you. Respondents categorize themselves instead of you categorizing them.
3. More questions can be asked because time is saved by the respondent simply checking.
4. Precoded answers are easily analyzed.

Disadvantages

1. It requires advance information about possible response categories that may be given.
2. It may bias responses by suggesting answers.
3. It does not allow for diversity and richness in individual expression.

You may decide to use both free-response and forced-choice questions to investigate the source area of interest. The free-response questions should be asked first to obtain the more general, less structured reactions of the respondent; forced-choice questions would be asked next to address those aspects of the issue that are particularly interesting to your study. In this way, a more balanced coverage of the area of interest can be made. The two types of questions tend to compliment each other when combined in this way. This technique of mixing item types is generally used when an interview is conducted using the questionnaire. It may or may not be of use in a mail survey.

Types of Forced-Choice Instruments

Some of the more commonly used types of forced-choice instruments are presented in this section. The four categories (check lists, inventories, ratings and rankings) presented indicate the basic range of instrument types, but only partially reflect the diversity of instruments that have actually been used. You may find it desirable to modify or combine these basic types to produce an instrument suited to your survey study.

A checklist is a series of items that are responded to by checking all that meet a certain condition. An example is:

> Have you played any of the team sports listed below during the last four weeks? If so, check all that you have played at least once during this time.
>
> ( ) baseball (hardball)
> ( ) softball (slow or fast pitch)
> ( ) football
> ( ) basketball
> ( ) volleyball
> ( ) soccer
> ( ) other team sport (please specify) _____
> ( ) none

Note that in the above example the respondent is not limited to choosing just one response. Following is a checklist that is used to identify where in a series of steps the respondent is located.

> Below are the requirements for graduation from the doctoral program. Check all that you have successfully completed to date.
>
> 1. Select an advisor . . . . . ( )
> 2. Applicancy . . . . . ( )
> 3. Select a Doctoral Advisory Committee . . . . . ( )
> 4. Course of study approved . . . . . ( )
> 5. Research or tool of language completed . . . . ( )
> 6. Comprehensive examination passed . . . . . ( )
> 7. Dissertation prospectus approved . . . . . ( )
> 8. Candidacy . . . . . ( )
> 9. Residency requirement met . . . . . ( )
> 10. Dissertation accepted . . . . . ( )
> 11. Oral Exam completed . . . . . ( )

Checklists may be analyzed differently for different purposes and to answer different questions. For example, the above team sports checklist could be analyzed by summing the number of checks for an individual respondent for playing and using the resulting number as as index of team sport participation or summing across individuals on football to get an index for the population. Similar checklists have been constructed for measuring acquisitiveness (how many of the following items do you own?), social economic level, inclination for self-medication, etc.

The graduation requirement checklist example could be used to monitor progress in the program. By combining the number of requirements passed with time enrolled in the program, it is possible to identify the average and the exceptional students as they progress through the program to identify those problem requirements that are not being met by many students. Although this example is relatively simple, this same principle may be used to monitor some more complex and diverse processes such as the implementation of an educational program in a school district.

While checklists only require respondents to provide yes/no data, an inventory requires respondents to additionally name or enumerate items or events that meet a certain condition. An example of an inventory item is:

> Approximately what percent of your total professional time this year are you spending in each of the following areas? (The total sum should equal 100%)
>
> | | | | |
> |---|---|---|---|
> | Teaching | ____% | Other Administration | ____% |
> | Research | ____% | Service | ____% |
> | Research Administration | ____% | Other (specify) | ____% |
>
> _____

Rankings and ratings are discussed together here because they both involve making quantitative judgments about some group of objects, persons,

activities, or ideas. Respondents make these quantitative judgments along some specified quality dimension which all of the things being judged have in common. This dimension may be, for example, monetary worth, utility for a given purpose, similarity to one's own feelings, frequency of use, or any of a number of other attributes.

Ranking requires a respondent to put a list of things in order from greatest to least, high to low, or most positive to most negative along some dimension. The things may be only partially ordered (of these five candidates for mayor, indicate your voting preference by putting a '1' next to your first choice, a '2' next to your second choice, and so on for all five candidates). The end results of ranking is an ordered list of things; however, this ordering does indicate how much better one thing is than others in the list. Also, information from a ranking does not indicate necessarily that the highest ranked thing is most desirable in any set of choices or good, but only that it is the best among the choices available. The data from this technique are ordinal in nature and it is important to match the analysis with the data. Analysis procedures appropriate for various measurement metrics are discussed later in this document.

Ratings involves comparing things on some scale which usually has from three to seven points. It is common to assume that these scales yield interval measures although theoretically they produce ordinal. (The interval scale of measurement allows for the use of an additional set of statistics). These ratings can be interpreted with reference to some criteria that is independent of the group of things being rated. That is, it is possible to say that something is "good", while with ranking it is only possible to say that something is "better than the other things". Examples of rating scale items are given on the next page.

Sample 1

The most important factor in determining job retention should be how well the person performs his/her duties.

( ) strongly disagree
( ) disagree
( ) neutral
( ) agree
( ) strongly agree

Sample 2

Directions: Read each sentence carefully and decide how often it is done in your class. Put one of the following letters in front of each statement which best describes your teacher's behavior.

Mark "A" if your teacher almost never does it in your class.
Mark "B" if your teacher seldom does it (less than ½ the time).
Mark "C" if your teacher sometimes does it (around ½ the time).
Mark "D" if your teacher often does it (more than ½ the time).
Mark "E" if your teacher almost always does it in your class.

1. _____ Our teacher tells us how to do a problem.

2. _____ In our class, the teacher gives a direct answer to our questions.

However, while the advantage of rating data is that each item is rated uniquely, a disadvantage is that all items may be rated high. Thus, the data might show that everything is "good" and you won't be able to show which is "best". That sort of conclusion comes from ranking items.

Selected Readings.

Babbie, E.R. Survey research methods. Belmont, Ca.: Wadsworth, 1973.

Chapter 7 discusses conceptualization and instrument design. Chapter 11 treats pre-tests and pilot studies of questionnaires.

Berdie, D.R., & Anderson, J.F. Questionnaires: Design and Use. Metuchen, J.J.: Scarecrow Press, 1974.

Chapter 4 and 5 discuss the design of the form and such support documents and procedures as cover letters, pre-letters

and follow-up on non-responses. The appendices include a number of excellent examples of questionnaires and accompanying letters.

Payne, S.L. The art of asking questions. Princeton, N.J.: Princeton University Press, 1951.

This is a clearly presented text on the types and pitfalls of question construction.

Selltix, C., Johoda, M., Deutsch, M., & Cook, S.W. Research methods in social relations. New York: Holt, Reinhart, & Winston, 1959.

Chapters 7 and 10 and Appendix C discuss questionnaire construction. This is a systematic presentation of basic principles of questionnaire construction and use.

Attitude Measures

Many times a major segment of the questionnaire is devoted to the measure of an attitude toward something. In these instances, the surveyor plans on adding several item values together to get a score for each respondent. In many ways the cluster of items which are designed to measure attitudes are an instrument within an instrument. Special care must be taken in the design and selection of these items so that the attitude score is meaningful.

Before discussing particular attitude scaling techniques, it may be appropriate to have a brief overview of the generic measurement scales distinctions which will be part of the concern when constructing any kind of items.

The three scales of measurement that need to be differentiated are nominal, ordinal and interval. The levels described briefly are:

Nominal. When responses are measured using a nominal scale, they are simply sorted into mutually exclusive categories. The only interpretations you can reach are that responses either are or are not in the same categories. Nothing can be concluded about whether one response is more or less than another. Examples are sex, major fields, political affiliation, etc.

Ordinal. With an ordinal scale, responses are categorized on the basis of relative standing. You cannot only discern whether responses are in the same category but also the categories allow you to compare responses as to whether they contain more or less of a property than another. Categories can be arranged sequentially acording to the amount of the property contained, but there is no guarantee that adjacent cate-

gories are the same distance apart as any other two categories. Examples are ranks of preference or descriptions like good, better, and best.

Interval. Interval data maintains the same comparison property as ordinal data with the exception that adjacent categories are some predetermined equal distance apart. Equal intervals between each unit is the characteristic that defines this scale. One of the most commonly used examples of an interval scale is temperature measured in Fahrenheit. Scores from achievement or attitude tests are often considered measured on an interval scale, but you will find some controversy as to the accuracy of this assumption.

In the area of attitude scale construction, four scaling procedures have received predominent attention: Likert, Thrustone, Semantic Differential, and Guttman. Although various assertions have been made about the relative advantages of one method over the others, the question of which technique is "best" is still largely unresolved. The particular requirements of the study, and your judgment and experience are still key factors in selecting the scaling procedure most appropriate for your study.

Likert scales. Likert (1932) scales typically contain 20 or more statements of belief which are either positively or negatively worded. A person responds to a Likert scale by indicating whether they strongly agree, agree, are uncertain, disagree, or strongly disagree with each statement. For example:

Example One

Teachers are nice people.

___strongly agree ___agree ___uncertain ___disagree ___strongly disagree

Example Two

I feel bad about my science class.

___always ___usually ___sometimes ___not often ___never

For each statement, values are assigned to these categories, usually 5, 4, 3, 2, and 1 respectively for positively worded statements and the reverse for negatively worded statements. (Zero is generally avoided because of possible problem with computers reading zero as missing information.) The responses for the items are summed, resulting in a single total score for the instrument. Note that the two examples different words were used to describe the five points on the scale. The use of these different words implies different scales of measurement. Thus, the responds to the two sample items would not be added together. Two common pitfalls of this type of measure are added together different scales and not making sure that a value of five represents the most positive response. In the two examples above five would be assigned to a response in the left most and right most response categories respectively. A 25 item scale would then have a possible range of 25 to 125, with the higher scores indicating more positive attitudes. Likert scores are interpreted by comparing scores against norms derived from populations of interest or by comparing individuals or groups to one another. While the data are generally seen to be ordinal in nature, an assumption of an interval scale is generally made for all analyses.

The general procedure for developing a Likert scale is outlined below. For greater detail on actually constructing such a scale, see Likert (1932).

1. Define the attitude variable and the population of people of interest.

2. Collect a wide variety of positive and negative opinion statements about the attitude variable, avoiding neutral statements.

3. Edit these statements into short, clear, simple statements of opinion.

4. Administer these items to a sample of the target population (about one to two hundred respondents are needed). Respondents rate each statement on a five-category rating scale: strongly approve, approve, undecided, disapprove, and strongly disapprove. These categories are assigned the values 5, 4, 3, 2, and 1 respectively. Negatively worded statements are reversed in the values assigned when scoring.

5. Items are selected for the scale by correlating item scores with total score. Items which correlate highly with the total score (i.e., those that discriminate well between high positive and low negative attitudes) are selected for the final scale.

6. The final scale is standardized by administering it to a representative sample of the target population. The norms thus derived are used in interpreting Likert scale scores.

Thurstone scales. Thurstone (1928) scales typically contain 20 to 30 statements of belief which vary from very negative to very positive statements, including items that reflect a neutral attitude. The items which appear on the instrument are selected by having a relatively large number of people judge the relative positiveness or negativeness of the statements. Data from the judges are then analyzed to select out those items which best represent the range of all possible attitudes with respect to a certain object. The items should have enough range to cover the scale, yet be judged unambiguous. While the development of the instrument is accomplished by two different stages, once developed the scored of an individual can be interpreted directly with respect to positiveness or negativeness and does not require a comparison group. "Scale values" are assigned to each item. The values ranges from 0, which is very negative, to 10, which is very positive. A person responds to a Thrustone scale by indicating agreement with the few statements

which are most similar to his/her beliefs. The average scale value of those statements agreed with is that person's attitude score. A brief example of Thurstone scale is:

> Directions: Please put a check ( ) in front of each statement with which you completely agree. If you disagree with a statement in any way, do not put any mark in front of it.
>
> ( ) 1. College develops self-reliance.
> ( ) 2. College has its drawbacks, but I like to go.
> ( ) 3. College is a waste of time.

For more comprehensive examples the reader is directed to Shaw and Wright (1967).

The following steps are used to develop a Thurstone scale. See Thurstone (1928) for greater detail.

1. Define the attitude variable.
2. Collect a wide variety of opinion statements about the specified attitude variable. Be sure to get statements that are fairly evenly graded from positive to neutral to negative statements.
3. Edit the statements into about 100 short, clear, simple statements of opinion.
4. Have between 50 and 300 (Thurstone suggests 200) judges sort the statements into 11 equal appearing intervals with 0 being most negative, 5 being neutral and 10 being positive.
5. Calculate the scale value of each statement.
6. Eliminate statements based upon the criteria of ambiguity and irrelevance.
7. Select about 20 statements with scale values evenly distributed along the scale. These items are then used to construct the instrument.

Semantic Differential. This scaling procedure (Osgood, Suci, & Tannenbaum, 1957) uses bipolar scales with contrasting adjectives at each end. Like the game of "Twenty Questions" this technique tries to identify a person's attitudes toward a concept by asking questions like:

"Is it interesting or boring?", "Is it near or far?", "Is it active or passive?". Respondents indicate their answers on a seven point scale interposed between the adjectives.

An example of a Semantic Differential scale would be:

> For each pair of words, please put an "X" on the line which best shows your feelings about educational evaluation. Don't skip any pairs of words, Use your first impression in marking each pair. Thank you for your cooperation.
>
> Education Evaluation
>
> Interesting_____ boring
> extremely quite slightly neutral slightly quite extremely
>
> far_____ near
> extremely quite slightly neutral slightly quite extremely
>
> active_____ passive
> extremely quite slightly neutral slightly quite extremely

The scales are scored by assigning 1 to 7 to the positions for each item, with 7 being at the most positive end.

Research has demonstrated that three main underlying factors account for most of the variablity of the ratings on the different adjective pairs. The three factors have been labeled Evaluation, Potency, and Activity. (See Heise, 1970, for a review of the research.) In combination these three factors provide a measure of a person's attitude toward a concept. A person's score on this type of instrument is interpreted by comparing it to some known distribution of scores.

The following general steps are involved in constructing a Semantic Differential Instrument. For additional information see Heise (1970).

1. Define the attitude variable.
2. Identify research on the factoral analysis of adjective pairs. Some collections of adjective pairs are reported in Osgood, et al., pp. 53-61 (1957), DiVesta (1966), Jakovits (1966), and Snider and Osgood (1968). By searching the literature you will probably be able to identify additional reports which are more appropriate for your particular study.

3. Select adjective pairs according to the following criteria:

   a. Select four adjective pairs for each factor (Evaluation, Potency, Activity) to be measured. This will usually produce adequqate reliability for scores.

   b. Select adjective pairs which have been researched with variables and populations as similar as possible to those with which you will be working.

   c. When possible, select adjective pairs which have "face" validity for the variable you will be investigating. Both "weak-strong" and "short-tall" may be part of the Potency factor, but the first has more "face" validity for the variable "Oil Embargo."

4. Construct the instrument and administer it to a sample of the population of interest to obtain norms and technical quality indices for interpreting Semantic Differential scores.

Guttman scales. Guttman (1944, 1947) scales are a fourth type of attitude scales. Because of the criteria used in constructing these scales, they tend to have fewer items than other scales, Guttman suggests ten. Items on a Guttman scale range from very negative to very positive along a single dimension. Although the statements all are phrased positively. An individual who responds positively to any particular item is assumed to also respond positively to all items of lower rank on that dimension, i.e., statements which represent a less positive attitude. The development of the instrument insures that the items are in rank order along a single dimension. An abbreviated example of a Guttman scale is:

____Yes ____No 1. An educated person has an elementary dipolma.

____Yes ____No 2. An educated person has a high school dipolma.

____Yes ____No 3. An educated person has a college dipolma.

Although a variety of item formats are used for these scales, items are usually scored as dichotomous items. If these items allow for more than a two option response types (yes/no, agree/disagree), then a

scoring rule is established to produce two option scores. For example, a Likert type item might be scored so that strongly approve or approve are a positive response and uncertain, disapprove, or strongly disapprove are a negative response. The aforementioned procedure gives Guttman scales the property of testing for "reproductability." This means that knowing the total number of positive responses a person gives allows you to reproduce the pattern of items on which the person gave positive responses. Reproductability becomes a criteria for scale construction. This is different than other scales, such as the Likert scales. On a Likert scale, marking half of the items "strongly approve" and the other half "strongly disapprove" will give the same score as marking them all "uncertain." In this sense, Guttman scales are easier to interpret than other scales. These scales are interpreted through norm group information. If scales that are already developed and are to be used with a new population, then the reproductability of the scale needs to be checked for the new population and new norm group distributions need to be developed.

The general steps for developing Guttman scales are presented below: For more information, see Guttman (1944, 1947) and Dotson and Summers (1970).

1. Define the attitude variable and population of people of interest.
2. Construct about 12 attitude statements which cover the attitude variable of interest.
3. Field test these items on about 100 people from the population of interest.
4. Statistically examine the field test responses to see if they form a Guttman scale. (A computer package to do this is now available in SPSS.) Reject items which do not fit the specified criteria.

5. Standardize the administration and gather normative data for the remaining items on a second sample of the population of interest. Between several hundred and 3,000 people are needed in this sample. The norms thus derived are used in interpreting Guttman scale scores.

To help you identify the scaling procedure most useful for your study, Table 1 describes these four most commonly used scaling procedures in terms of the measurement metric obtained from each scale type, the relative difficulty of constructing scales according to each procedure, and one or two outstanding characteristics of each procedure.

Table 1

Comparison of Four Scaling Procedures[a]

Scaling Procedure	Measurement Metric	Difficulty of Construction	Outstanding Characteristics
Likert	ordinal	medium	Likert-like item formats are the most commonly used in attitude scales, thus respondents are likely to be familiar with the format.
Thurstone	interval	hard	Because of judges' ratings, scores have a criterion as well as normative interpretation.
Semantic Differential	ordinal	easy	a) The same scales can be used to measure different attitude variables. b) Three scores are available for each attitude variable.
Guttman	ordinal	hard	a) Scales developed by this procedure tend to be unidimensional. b) To a large degree, the pattern of responses on individual items can be reproduced from the total score.

[a]See Fishbein (1967) for similar table and elaborated discussion.

Selected Readings.

Fishbein, M. (Ed.) Readings in attitude theory and measurement. New York: John Wiley, 1967.

Summers, G.F. (Ed.) Attitude measurement. Chicago, Rand McNally, 1970.

Both of these books of readings contain classic articles on attitude measurement. The Fishbein book emphasizes attitude theory; the Summers' book emphasizes attitude measurement techniques.

Lake, D.G., Miles, M.B., & Earle, R.B., Jr. Measuring human behavior: Tools for the assessment of social functioning. New York: Teachers College Press, 1973.

Shaw, M.E., & Wright, J.M. Scales for the measurement of attitudes. New York: McGraw-Hill, 1967.

The Shaw and Wright book contains 176 attitude scales from research reports and fugitive sources. The Lake, Miles, and Earle book reviews 84 instruments dealing with social functioning; it also contains a review of 20 other collections of instruments or reviews of instruments for social research. Both books are excellent sources for instruments, or starts on instruments if you are developing your own.

Miller, D.C. Handbook of research design and social measurement. (2nd ed.) New York: David McKay, 1970.

The first three parts of this book provide guides on research in the social sciences; the fourth part contains reviews and examples of 33 scales or indexes for research; and the fifth part deals with funding and reporting research. Pages 88 through 99 deal particularly with the selection and construction of attitude scales.

III. SURVEY MANAGEMENT

The Mechanics

This section of the document refers solely to the mechanics of mail surveys. While some of these factors may also be considerations when data are gathered by handing out the instruments to intace groups (such as a classroom) or by interviews whether field or phone; other "red flags" are specific to those techniques and not mentioned in this section. Individuals who are interested in the problems of data collection and handling in those other forms should refer to other documents.

Delivery of the questionnaire. Once you are satisfied with the format and content of the questionnaire, you are ready to distribute it to the sample. It is the dream of every investigator that 100% of the questionnaires sent out will be returned completed and ready to go for analysis. One hundred percent response is not often seen in the survey literature. How close one comes to it depends on many criteria, but here is a list of questions that may help you in identifying some of the factors which influence response rate.

Some Factors Effecting Response Rates[a]

1. Are there sufficient funds for a follow-up, and has the follow-up procedure been planned?
2. Is the questionnaire or cover letter too folksy or too formal for the respondents?
3. Has respondent anonymity or confidentiality been guaranteed?
4. Has the sponsorship and rationale for the survey been identified?
5. Is the questionnaire aesthetically pleasing?

6. Is the questionnaire too difficult or time consuming to complete?
7. What kind of incentive or motivation has been provided for completing the questionnaire?
8. Is the mailing list being used current and correct?
9. Does the questionnaire appear as part of a bulk mailing, or is there a personalized touch?
10. Will the questionnaire be delivered at a business or home address?

[a]From Berdie and Anderson

There are many creative ways that you can increase the response rate. For instance, one company, in an effort to get important information from doctors, printed the questionnaire on the back of a check. The endorsed check with completed (hopefully) questionnaire was then returned through the banking system.

After the Delivery. After the questionnaires have been sent, there is a period of limbo before they begin to return. This period of time can successfully be used to prepare for later aggregating, analyzing, and reporting.

Each day as the questionnaires return, there should be a record of the number returned that day. This daily return count can be useful for identifying the point at which follow-up letters should be sent. This record will also be a useful reference later for anyone, including yourself, as an indication of the necessary time between the first mailing and the subsequent follow-ups.

While keeping a daily count, it is also useful to place an identifying code on each questionnaire as it is returned. There are two ways this can be done. One procedure is to identify each questionnaire by sequentially numbering them as they return. If the questionnaires do not already have an individual identifying code, this number should be used

as such. If the questionnaires already have individual identifying numbers, the date of return could be used for all questionnaires returned on a particular day. The purpose of this numbering is for differentiating early respondents from late respondents. The question may be raised of whether early respondents differ from late respondents, or respondents receiving a follow-up differ from those who didn't. If you haven't planned ahead, these questions will be unanswerable.

In the final report, thorough readers will want to know the information contained in the "tree" diagram in Figure 5. This information allows them to judge the credibility of the survey results.

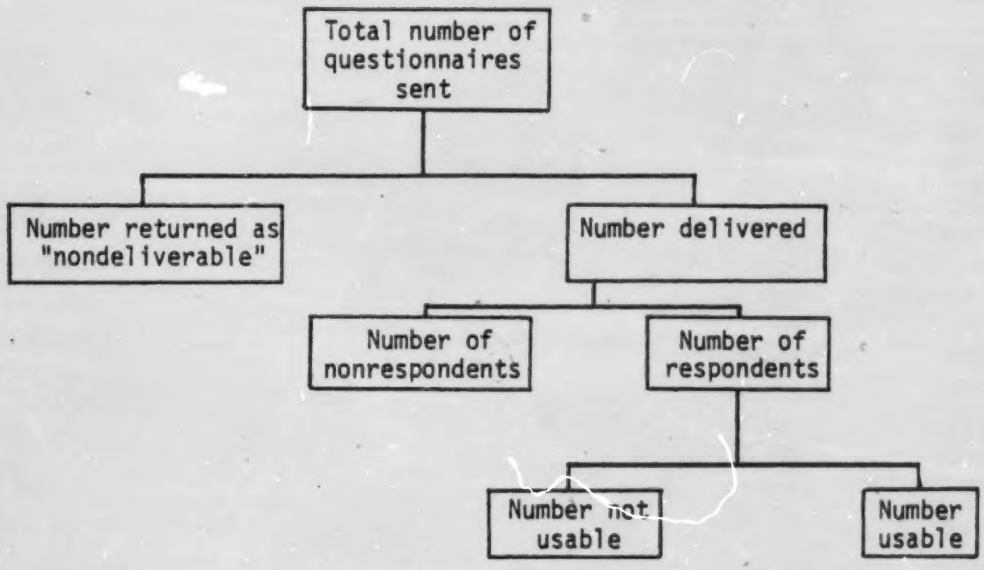

Figure 5

Keeping Record for Calculation of the Response Rate

Notice that the left box in each "branch" of the "tree" represents missing information. The individual who is collecting the survey information

should provide the reader of the report with reasons for categorizing forms in the left box at each level of the "tree". In some cases, this will require special effort on the part of the survey conductor.

Coding. Coding is the act of classifying the diversity of individual responses to items in the questionnaire into a limited number of categories.

For closed-end items, the categories have been chosen before the administration of the instrument. Coding in this case is a comparitively straightforward process. Numbers are assigned to each of the possible categories. The concern is whether the numbers assigned are to reflect quantitatively some qualitative characteristics in the data. On one hand, the numbers assigned can simply differentiate categories or, on the other hand, they can contain information on magnitude. In other words, the scale of measurement associated with the data must be considered. Failing to make this difference during coding is not fatal. If it is found later that a different coding is needed, most computer programs allow for a transformation of the codes. It could also be done by hand. It is best, however, not to have to do either; it is just not an efficient use of time. The meaningfulness of the number is indicative of the measurement metric and has consequences for the analysis techniques which can appropriately be used.

For open-ended (or free response) items in a questionnaire, the coding process is somewhat more involved. You could decide the categories beforehand and the coding becomes like that involved with closed-end items. Typically, the categories are not decided upon until the questionnaires have been returned and there has been an opportunity to review the responses.

Content analysis consists of the following steps:

1. Pick a random sample of questionnaires (around 100 if possible).

2. Read the responses to the item and sort them into categories that are descriptive of the response.

3. Decide if the categories reflect the range of responses adequately. If not, reorganize.

4. The result should be a workable number of categories that are indicative of the response range. Each category should clearly not overlap with other categories, while taken as a whole all categories cover the response range.

5. Problems often encountered are:

    a. An "other" category that contains a large proportion of the responses.

    b. Categories so broad that all responses fall into one category or a response fits into more than one.

    c. Categories so specific that the coder finishes with as many categories as responses.

Once the categories have been described and codes assigned, coders will then read the entire batch of returned questionnaires and decide on the basis of this category system what code should be assigned. Provision should be made for the addition of new categories if important ones arise during the coding. It is a headache to add a new code midstream, (all questionnaires have to be coded again) so it should be avoided unless the response category adds considerably to the analysis.

There is nothing inherently sacred about using numbers as code categories. Alphabetic codes as A, B, or C should also be assigned. Be aware, though, that if a computer program package is to be used to analyze the results, some programs are not capable of handling anything but numbers. Also, from a keypuncher's view, it is easier to consistently use numbers. If keypuncher's have to switch between alpha and numeric systems, they are more likely to make errors.

In the literature, you may come across a differentiation between "precoding" and "coding." "Precoding" refers to assigning codes before the administration of the instrument and "coding" to after. Just because an item calls for free response from the respondent, doesn't mean that you must wait for the responses to think about coding schemes. Coding anticipation may identify ambiguous items. For example, a question like "How often, on the average, do you ride a bicycle each week?" could elicit possible responses of "very often," "½ hour a day," or "2 days a week"--a coding headache. Either the question must be written more specifically or limits placed on the form of the response, i.e., How many hours a week, on the average, do you ride a bicycle?

It should be evident that the coding scheme used in a study depends on the questions asked and the particular way the responses will be analyzed. There are, however, some coding suggestions that can be made:

1. Create categories that are exhaustive of the response range but mutually exclusive of one another.
2. Be consistent in your coding across items, such as always using "0" for "no" and "1" for "yes." Don't code "1" for "no" on different questions.
3. Maintain detail; remember that you can always collapse categories in analysis but it is difficult to sub-divide.
4. Don't reinvent the wheel if preestablished coding schemes exist for your items. (An example for occupations is in Oppenheim, 1966, pp. 263-284.)
5. Avoid blanks and alphabetic characters because of the idiosyncracies of some computer program. For instance, in SPSS (Nie, et al., 1965) blanks are treated as zeros. Consequently,

if you allow blanks and have zero as a code your data will be difficult to computer analyze.

6. If possible, differentiate missing data. Distinguishing between the possible answers below may be important in the analysis.

"Don't know",	Code = 1
"Skipped the question",	Code = 12
"Refused to respond",	Code = 13
"Not decipherable",	Code = 14
"Inappropriate response", etc.	Code = 15

7. Consider items that allow multiple responses to see if a unique coding system would be of use (See Muehl, 1969, p. 17 for an example of using a geometric code for such items.)

The coding process. There are many possible procedures available for getting data from the questionnaires to computer usable data. The one chosen will depend on the design of your questionnaire and the resources available. Figure 6 shows some of the possible paths you might use. Keep in mind the tradeoffs between time involved in keypunching or coding, costs, and the limitations placed on the respondent and the responses.

Path 1-2-3a-4-5. This is the most commonly used procedure. It allows for the coding of both open-end and closed-end items. The main drawback is the number of persons (coders and keypunchers) that must be involved, thus increasing the probability of clerical or transfer errors.

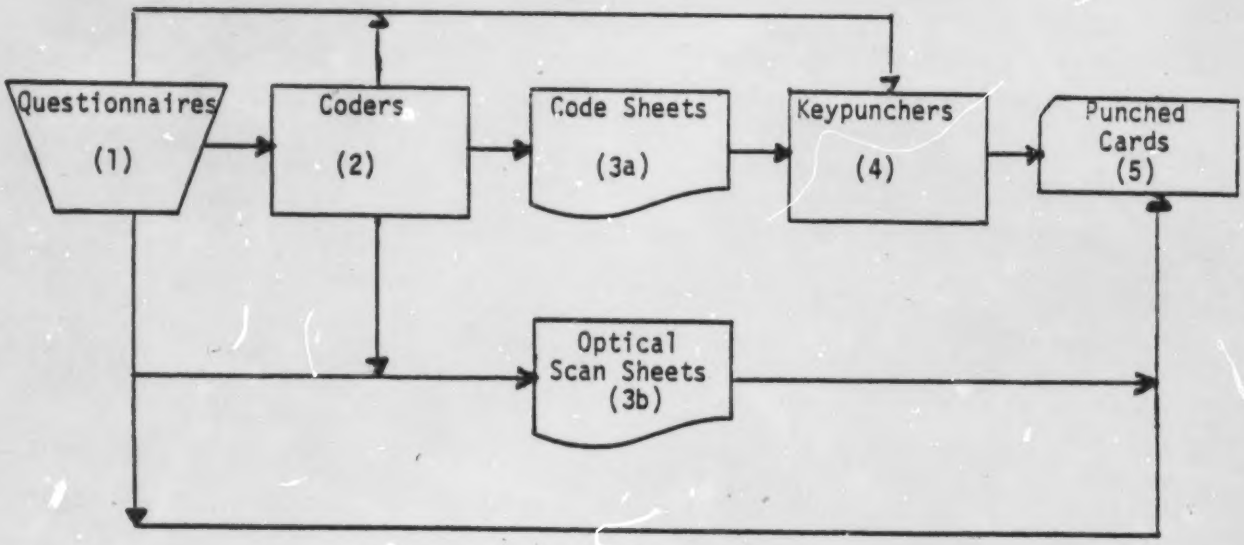

Figure 6

The Coding Process

Path 1-2-4-5. This process involves a procedure called "edge coding" where coders write the code on the edge of the questionnaire. The keypuncher then works directly from the questionnaire. It is helpful to indicate on the questionnaire the column(s) that will be used for the code so that the keypuncher does not inadvertantly make an error.

Path 1-2-3b-5. Coding is done onto optical scan sheets that are read by an optical scanner and punched cards or type are produced. This eliminates the need for keypunching and thus a possible source of error, while introducing errors in tranferring onto the optical scan sheets.

Path 1-4-5. This procedure could be used if all items are closed end and codes are obvious to the keypunchers.

Path 1-3b-5. This requires the respondent to mark the responses on a separate optical scan sheet. Only closed-end items can be used.

Path 1-5. The questionnaire is printed on a sheet that will be optically scanned. Closed-end items are used and the length of the questionnaire is usually one page. Special printing of the questionnaire is required, and you must be confident that respondents will follow directions.

Documentation of the coding frame, the set of rules for classifying the responses to each item, is essential. The product of the documentation process is the codebook, the foundation of the coding processes.

This codebook must contain the following details:

1. General information.
    a. Project name and identification number
    b. Computer file names where data are stored
    c. Location of data decks or magnetic tapes where data will be archived.
2. A copy of the questionnaire that is being described
3. The coding frame for each item. A suggested layout of columns of each page in this section is:

Column(s) Designation	Item or Question Number	Variable Name	Description of Variable	Assignment Categories
77-78	12	SCCSS	View of success of program	01=Very unsuccessful 02=Unsuccessful 03=Undecided 04=Successful 05=Very successful 08=Unusable response 09=Unanswered 10=Two responses given

Having each coding scheme on a separate page in the codebook may alleviate the possible problem of coders using the wrong system for some particular question or variable.

4. Description of any codes that are not specific to items on the questionnaire. Make sure that any column on a punched card or tape, for example, has an explanation. While you think you will remember why there are "1's" in column 34 for some respondents, you really won't.

5. Any additional information. This includes information that may clarify the coding process such as a list of coding problems that occurred. This will be of help if a similar coding process or questionnaire is again considered.

The ultimate rule in the construction of the codebook is to record absolutely everything; trust nothing to memory or to obviousness. There is no denying the fact that the construction of a codebook is a dull and enervating job. The result, however, is a valuable document to anyone involved with the development of the report such as programmers, coders, and report writers. External to the project, the codebook will also be helpful to researchers who are considering similar projects, secondary analysts (See Cook, 1974 for an explanation of this area), or meta-analysts (See Glass, 1976).

Data cleaning or error checks. Unfortunately you can never rest assured that all the information starting from the returned questionnaires through to the final analysis is error-free. Errors can occur with any person from the respondent incorrectly marking a question to a secretary typing the final report. Figure 7 (from Tull and Albaum, 1973) is a good illustration of the many points in a survey where errors could occur. There are, however, precautions and procedures that can be used to ease this fear of faulty data. The extent to which you go into error checking ("data cleaning") depends on the accuracy necessary for the data, the importance of the decisions to be made on the basis of the data, and the amount of time and money available for the error checks.

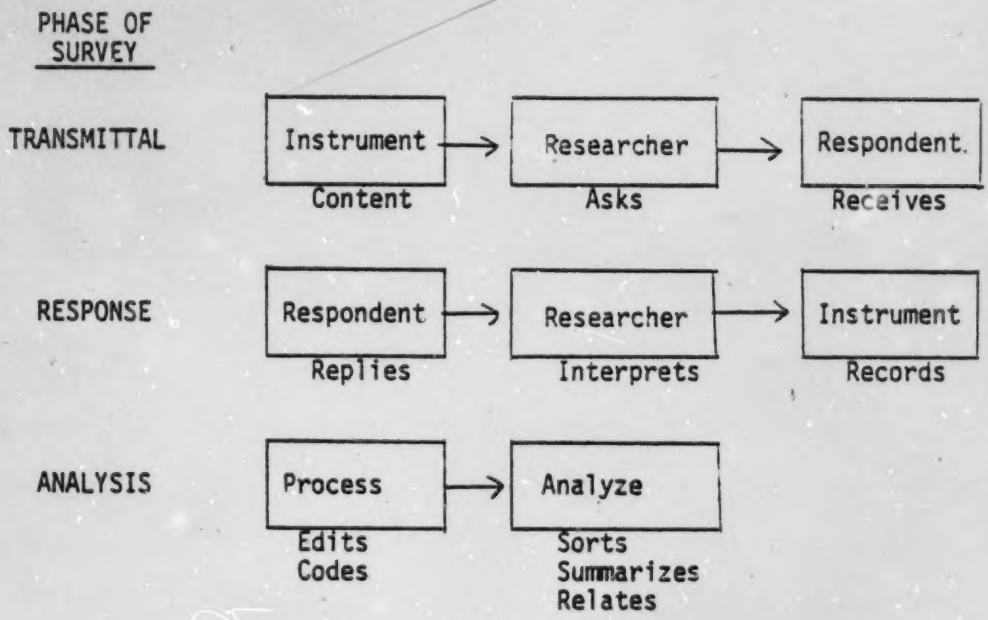

Figure 7

Potential Sources of Error in Sample Surveys[a]

[a]from Tull and Albaum, 1973, p. 62.

One place that errors occur is during the coding process. One technique to avoid error, or at least to pinpoint an error-abundant coder, is for a supervisor to be assigned to randomly check the work that coders have done. Also coders should cross check samples of each other's work on a given schedule.

Still another source of error is the step from code sheet to keypunched cards or tape. One method often used to check for keypunching errors is "verifying." This involves rekeypunching the already punched cards. The keypunching machine notes any discrepancies between what is punched on the card and what is keypunched the second time. There is some question as to the benefit of verfying (see Nemanich, 1972).

Since it involves a complete repunching, the process of verifying takes twice as long and it is twice as expensive.

Another possible means of data cleaning is a card sorter. The card sorter can be set to sort cards as to the punches in a particular column. Cards with illegal punches in a column can be identified and checked. Unfortunately, this procedure will not identify cards with the incorrect but legal code in a column.

You can also employ "contingency checking". Here, in a preanalysis, certain variables are crossed (as in a contingency table) with other variables to identify suspect data. For example, this could identify the highly unlikely situation of an eight-year-old with a Ph.D. This procedure will only identify errors that are extreme. Other "possible" errors may slip through.

Selected Readings.

Babbie, E.R. Survey research methods. Belmont, Ca.: Wadsworth Publishing Company, 1973.

Chapter 7 deals with precoding and reproducing. Chapter 8 is concerned with mail distribution, monitoring returns, and acceptable response rates. Finally, Chapter 10 covers topics in data processing, coding, and data cleaning.

Backstrom, C.H., & Hursh, G.D. Survey research. Evanston, Ill: Northwestern University Press, 1963.

This text has several good checklists dealing with different segments of survey methodology. Chapter VI has checklists on contents of the coding manual and coding hints.

IV. DATA ANALYSIS

Nonresponse Analysis

If you obtain a 100% return on a questionnaire, the only errors to be concerned with are those from sampling. However, once returns are below this 100% level, there is always a question of how the people that did not respond would have affected the results had they sent in returns. Rules for acceptable response rates vary. Babbie (1973) suggests that a 50% return rate is the lowest that can be considered "adequate", but some researchers won't accept anything below 80%. In any case, there are varying ways of accomplishing an analysis of the possible differences between respondents and nonrespondents. Most involve contacting a sample of nonrespondents through a different method, i.e., phone or field interviews to get their responses, and then comparing these to the returned responses. Other techniques involve comparing demographic variable results (e.g., ages, sex) from the respondents to those known of the general target population. Any significant differences will lead to toning down the generalizability of the results. A match, however, will not guarantee representativeness.

Selected Readings.

Daniel. W. Nonresponse in sociological surveys. Sociological Methods and Research, 1975, 3, 291-307.

There have been many articles written concerning the biases involved in surveys because of nonrespondents. This article presents eight methods of nonresponse analysis with description and comments.

Using the Computer

Because of the availability of computer facilities and the decrease in cost of computer time, it is difficult to think of many situations when manual analysis is superior to computer. Such instances might be where the sample size is small, the number of operation is few, or budget is so tight so as to eliminate the computer as a viable option. It will be assumed for the rest of this document that a computer will be used for most of the data analysis.

How does one go about arranging to computer analyze the data? One method is to hire a professional computer programmer to write an analysis program and an output format specifically to your questionnaire data and analysis guidelines. This is the most expensive approach and should be used only when the analysis procedure is very unique or when a "canned" program (preprogrammed or packaged program) is not available. The best advice is to explore what's available at your site. There exist many canned programs such as SPSS (Nie, et al., 1975) that are very flexible and contain most analyses that will be needed. Besides these published programs, most larger universities, colleges, or school districts have their own local program packages that could be used.

The ease of using the computer leads to two warnings of which you should be aware. First, because the survey analyst can obtain hundreds of crosstabulations (for example, 30 items can yield 435 possible tables of just two-way crosstabulation), there is the temptation to become involved in grandiose data-snooping expeditions. Many surveys have exploration as their underlying purpose, but keep in mind the primary purpose of the study and perform exploration related directly to the study (see Selvin & Stuart, 1966). According to the laws of probability

if you search long enough, you will always find something statistically significant and there will probably be some way of interpreting it (see Wincks & Campbell, 1969). Do not become so rigid, though that you ignore all other information an do only pre-planned analysis (see Barber & Fox, 1968 for an elaboration on the values of serendipitous findings).

The second warning is to avoid wallowing in the sea of available statistical procedures offered by computer programs. As Einhorn (1972) warns, there is a temptation to substitute immediate empirical analysis for more analytic thought. Almost any statistical procedure can be applied to a group of data (see Lord, 1953), but it is your responsibility to understand the mathematical bases and underlying assumptions of a procedure, and whether your data set meets those assumptions.

Selected Readings.

Nie, N.H., et al., Statistical package for the social sciences. New York: McGraw-Hill, 1975.

This is a widely used canned program for analyzing social science data. It's a relatively easy program to understand and offers statistics ranging from means and standard deviations to factor analyses. A plus for this manual is that not only does it instruct the user how to obtain the statistics, but it also gives brief and clear explanation of the statistics, their possible uses, and further references.

Silvey, J. Deciphering data. New York: Longman, 1975.

This small book is helpful to the reader by suggesting ways to code data so as to avoid pitfalls with using the computer. There are also helpful suggestions dealing with constructing the codebook and coding the questionnaire.

Choosing Statistics

Textbooks often divide the use of statistics into two categories: descriptive and inferential. Descriptive statistics are used to describe a particular sample and inferential statistics are used to describe the population from which the sample was chosen. All statistics cannot automatically be categorized as one of the two, though. For instance, a mean can be calculated to describe a sample, but then it can also be used to infer about a population. Thus, the categories describe the use of the statistic and not the statistic *per se*.

Before you consider what statistics to use with the data, the measurement metric of the data must be realized. As previously mentioned in this document, one of the most frequent statistical transgressions is using statistics designed for one type of measurement metric when data were collected or recorded at a different level without acknowledging the possible misinterpretations.

A short description of the scales is presented here for review:

Nominal. When responses are measured using a nominal scale, they are simply sorted into mutually exclusive categories. The only interpretations you can reach are that responses either are or are not in the same categories. Nothing can be concluded about whether one response is more or less than another. Examples are sex, major fields, political affiliation, etc.

Ordinal. With an ordinal scale, responses are categorized on the basis of relative standing. You cannot only discern whether responses are in the same category but also the categories allow you to compare responses as to whether they contain more or less of a property than

another. Categories can be arranged sequentially according to the amount of the property contained, but there is no guarantee that adjacent categories are the same distance apart as any other two categories. Examples are ranks of preference descriptions like good, better, and best.

Interval. Interval data maintains the same comparison property as ordinal data with the exception that adjacent categories are some predetermined equal distance apart. Equal intervals between each unit is the characteristic that defines this scale. One of the most commonly used examples of an interval scale is temperature measured in Fahrenheit. Scores from achievement or attitude tests are often considered measured on an interval scale, but you will find some controversy as to the accuracy of this assumption.

A number of research, measurement and statistics texts will also discuss a measurement scale called ratio. In this scale an additional definition of the existence of a meaningful zero is added to the "equal ness" of intervals. An example of this scale might be the exact number of seconds it takes an individual to complete a task. However, since ratio scales are often more evident in the natural sciences than in the social sciences this scale is not part of the discussion of statistics. If you believe one of your measures is a ratio scale the following discussion will not assist you in choice of analysis.

To discuss all available statistical procedures would be a herculean task for this document. A display of the names of some of the more common techniques appears in Figures 8 and 9. Figure 8 presents some options for descriptive statistics and for describing the association between two variables. Figure 9 presents the names of some of the more common inferential techniques. Do not interpret methods that appear in the same cells

		Level of Measurement		
Type of Statistic	Level of Second Variable	Nominal	Ordinal	Interval
Central Tendency		Mode	Mode Median	Mean (Arithmetic, Geometric, Harmonic, Contraharmonic) Median Mode
Dispersion		Percent in the Mode Percents in the distribution categories	Percentile Differences Interquartile Range Range	Standard deviation (variance) Interquartile Range Percentile Differences Range
	Nominal	Chi Square Family Chi Square Contingency Coefficient Tschruprow's T Cramer's V Phi Goodman-Kruskal's Tau Lambda Tetrachoric Correlation Kappa Kim's Uncertainty Coefficient Epsilon Yule's Q		Point biserial correlation Biserial correlation Triserial correlation Eta
Association	Ordinal		Spearman's Rho (rank-order correlation) Goodman-Kruskal's Gamma Kendall's Tau Somer's D	
	Interval			Product Moment Correlation Correlation ratio, eta

Figure 8

Descriptive Statistics

		Level of Measurement					
		Nominal		Ordinal		Interval	
Statistical estimation techniques		Confidence interval for a proportion		Confidence interval for a proportion		Confidence interval for a mean	
Measures of significance	One-sample test	Chi Square goodness-of-fit test		Kolomogorov-Smirov test Runs test		z test (large sample) t test	
	Two-sample test	INDEPENDENT Chi square test of independence Fisher's exact test z test for differences between proportions	CORRELATED McNemar test	INDEPENDENT Wald-Wolfowitz runs test Mann-Whitney U test	CORRELATED Sign test Wilcoxon matched pairs-signed ranks test	INDEPENDENT Independent t test	CORRELATED Correlated groups t test
	K-sample test	Chi square test for K Independent samples	Cochran Q test	Kruskal-Wallis H test Median test	Friedman two-way analysis of variance test	F test (ANOVA)	

Figure 9

Inferential Statistics

as interchangeable. Assumptions and interpretations are sometimes entirely different. The figures are provided so that you can reasonably access a statistics text.

Not shown in the figures are statistics for studying relationships between three or more variables. If you are interested in these kinds of relationships, you should see literature about:

Partial and part correlations (Glass & Stanley, 1970)

Multiple regressions (Draper & Smith, 1966)

Discriminate Analysis (Cooley & Lohnes, 1978: Tasouka, 1970)

Path Analysis (Blalock, 1972)

Factor Analysis (Harman, 1977)

Multidimensional Scaling (Shepard, et al., 1972)

Cluster Analysis (Everitt, 1974)

These analyses methods are relatively sophisticated, be sure to read about their use before you ask the computer to output the results for you.

Measures of association between two variables. The most common investigation in surveys is looking at the association between two variables. If you are concerned with exploring for relationships with a sample of data, become familiar with Rosenberg's The Logic of Survey Analysis (1968). This text gives an excellent discussion of the types of relationships between two variables, the ways that these relationships may be hidden, and possible procedures for their discovery.

As for statistics to be used, do not limit yourself to the Pearson Product Moment Correlation (r), or the Phi Coefficient (ϕ), the two most used. Investigate the many other statistics, some of which may be more

appropriate. The survey literature abounds with studies that use η and φ; however, PRE (proportionate reduction in error) measures are not frequently seen. PRE statistics give an estimate of how much better you can predict information about one variable by knowing certain information about the other. PRE's are in many cases found to be more appropriate (see Costner, 1965).

Loether and McTavish (1976) state four characteristics that should be reported whenever a statistic for an association is used:

Existence — You should be able to state whether any kind of relationship exists or not.

Strength — If a relationship is found, a noting of its strength is essential. For example, a correlation of .10 can be found to be statistically significant, but it certainly cannot be said to be an indication of a strong relationship.

Direction — For statistics like correlations, you should note whether the relationship is negative or positive; for asymmetric PRE statistics, it is necessary to identify the independent variable and the dependent variable.

Nature — This refers to the type of relationship that the statistic describes. For instance, is it a statistic that describes curvilinear or linear relationships; prediction of categories or order of categories?

Selected Readings.

Andrews, F.M., Klem, L., Davidson, T.N., O'Malley, P.M., & Rodgers, W.L. A guide for selecting statistical techniques for analyzing social

science data. Ann Arbor, MI.: Institute for Social Research, 1974.

This document provides the reader with a decision-tree approach to selecting an appropriate statistic to use in the analysis of data. More than 100 different statistics and statistical techniques are presented with excellent references to their descriptions.

Huff, D. How to lie with statistics. New York: Norton, 1954.

This book has become a classic in warning the reader or user of statistics of the possible misinterpretations using seemingly logical methods of presentation.

Rosenberg, M. The logic of survey analysis. New York: Basic Books, 1968.

This is an excellent book dealing with how to use data to discover relationships among variables for interpretation of results. It avoids the pitfalls of data snooping by excellent sections describing kinds of relationships and warning of spurious relationships

Glass, G.V., & Stanley, J.C. Statistical methods in education and psychology. New Jersey: Prentice-Hall, 1970.

Loether, H.J., & McTavish, D.G. Descriptive and inferential statistics. Boston: Allyn and Bacon, 1976.

Siegel, S. Nonparametric methods for the behavioral sciences. New York: Free Press, 1968.

These three textbooks cover very well the different statistics that one will commonly find in survey literature. For the most part, foreboding derivations are avoided or at least forewarned. The Loether and McTavish book is excellent for descriptive statistics; Glass and Stanley for inferential statistics; and Siegel for non-parametrics.

PART II

Bibliography of Survey Methodology

1. Alderfer, C. Comparison of questionnaire responses with and without preceding interviews. Journal of Applied Psychology, 1968, 58, 335-340.

2. Allen, E.L. Techniques of attitude scale construction. New York: Appleton-Century-Crofts, 1957.

3. Alutto, J.A. Some dynamics of questionnaire completion and return among professional and managerial personnel: The relative impact of reception at work site or place of residence. Journal of Applied Psychology, 1970, 54, 430-432.

4. American association for Public Opinion Research. Code of professional ethics and practices. Public Opinion Quarterly, 1960, 24, 529-530.

5. American Psychological Association. Ethical principles in the conduct of research with human participants. Washington, D.C.: American Psychological Association, 1973.

6. Anderson, J.F. & Berdie, D.R. Effects on response rates of formal and informal questionnaire follow-up techniques. Journal of Applied Psychology, 1975, 60, 225-227.

7. Anderson, N.H. Scales and statistics: parametric and non-parametric. Psychological Bulletin, 1961, 58, 305-316.

8. Andrews, F.M.; Klem, L.; Davidson, T.N.; O'Malley, P.M.; & Rodgers, W.L. A guide for selecting statistical techniques for analyzing social science data. Ann Arbor, Michigan: Institute for Social Research, 1974.

9. Andreasen, A. Personalizing mail questionnaire correspondence. Public Opinion Quarterly, 1970, 34, 273-277.

10. Ash, P. & Abramson, E. The effect of anonymity on attitude-questionnaire response. Journal of Abnormal and Social Behavior, 1952, 47, 722-723.

11. Azrin, N.H.; Holz, W.; & Goldiamond, I. Response bias in questionnaire reports. Journal of Consulting Psychology, 1961, 25, 324-426.

12. Babbie, E.R. Survey research methods. Belmont, California: Wadsworth, 1973.

13. Bachrack, S.D. & Scoble, H.M. Mail questionnaire efficiency: Controlled reduction on nonresponse. Public Opinion Quarterly, 1967, 31, 264-271.

14. Backstrom, C.H. & Hursh, G.D. Survey research. Evanston, Illinois: Northwestern University Press, 1963.

15. Baier, D. E. Reply to Travers' "A critical review of the validity and rationale of the forced-choice technique." Psychological Bulletin, 1951, 48, 421-434.

16. Barber, B. & Fox, R.C. The case of the floppy-eared rabbits: An instance of serendipity gained and serendipity lost. American Journal of Sociology, 1958, 54, 128-398.

17. Barnette, W.L., Jr. The non-respondent problem in questionnaire research. Journal of Applied Psychology, 1951, 34, 397-398.

18. Bass, B.M.; Cascio, W.F.; & O'Connor, E.J. Magnitude estimations of expressions of frequency and amount, Journal of Applied Psychology, 1974, 59, 313-320.

19. Bauer, R.D. & Meissner, F. Structures of mail questionnaires: Test of alternatives. Public Opinion Quarterly, 1963, 34, 397-398.

20. Baur, E.J. Response bias in a mail survey. Public Opinion Quarterly, 1947, 11, 594-600.

21. Belson, W.A. & Duncan, J.A. A comparison of the check-list and open-response questionnaire system. Applied Statistics, 1962, 11, 120-132.

22. Bender, D.H. Colored stationery in direct-mail advertising. Journal of Applied Psychology, 1957, 41, 161-164.

23. Bendig, A.W. Reliability and the number of rating scale categories. Journal of Applied Psychology, 1954, 38, 38-40.

24. Bennett, E.M.; Blomquist, R.L.; & Goldstein, A.C. Response stability in limited-response questioning. Public Opinion Quarterly, 1954, 10, 234-241.

25. Benson, L.E. Mail surveys can be valuable. Public Opinion Quarterly, 1946, 10, 234-241.

26. Berdie, D. Mail questionnaire response rates: Updating outmoded thinking. Journal of Marketing, 1976, 40, 71-73.

27. Berdie, D.R. & Anderson, J.F. Questionnaires: Design and use. Metuchen, New Jersey: Scarecrow Press, 1974.

28. Berdie, D.R. Questionnaire length and response rate. Journal of Applied Psychology, 1973, 57, 278-280.

29. Berdie, F. What test questions are likely to offend the general public? Journal of Educational Measurement, 1971, 8, 87-93.

30. Berger, P.K. & Sullivan, J.E. Instructional set, interview context, and the incidence of "don't know" responses. Journal of Applied Psychology, 1970, 54, 414-416.

31. Bevis, J.C. Economical incentive used for mail questionnaire. Public Opinion Quarterly, 1948, 12, 492-493.

32. Black, J.A. & Champion, D.J. Methods and issues in social research. New York: John Wiley & Sons, 1976.

33. Blalock, H.M., Jr. Casual inference in nonexperimental research. Chapel Hill: University of North Carolina Press, 1964.

34. Blalock, H.M., Jr. Social statistics. (2nd ed.) New York: McGraw-Hill, 1972.

35. Blankenship, A.B. The choice of words in poll questions. Sociology and Social Research. 1940, (September - October), 12-18.

36. Blankenship, A.B. Consumer and opinion research: The questionnaire technique. New York: Harper & Bros., 1943.

37. Bock, R.D. & Jones, V.L. The measurement and predictions of judgment and choice. San Francisco: Holden-Day, 1968.

38. Boek, W.E. & Lade, J.H. A test of the usefulness of the postcard technique in a mail questionnaire study. Public Opinion Quarterly, 1963, 27, 303-306.

39. Bogart, L. No opinion, don't know, and maybe no answer. Public Opinion Quarterly, 1967, 31, 331-345.

40. Boruch, R. F. Maintaining confidentiality of data in educational research: A systemic analysis. American Psychologist, 1971, 26, 413-430.

41. Boruch, R.F. Relations among statistical methods for assuring confidentiality of social research data. Social Science Research, 1972, 1, 403-414.

42. Boruch, R.F. & Cecil, J.S. Methods for assuring the confidentiality of social research data. Book in preparation, 1976.

43. Bracht, G.H. & Glass, G.V. The external validity of experiments. American Educational Research Journal, 1968, 5, 437-474.

44. Bradt, K. The usefulness of a post card technique in a mail questionnaire study. Public Opinion Quarterly, 1955, 19, 218-222.

45 Brandt, L. A neglected issue concerning nonrespondents to mail questionnaires. International Mental Health Research Newsletter, 1965, 8(2), 12-16.

46. Brenan, R.D. Trading stamps as an incentive in mail surveys. Journal of Marketing, 1958, 22, 306-307.

47. Brooks, V. Can you trust mail questionnaires? Printers' Ink, 1947.

48. Brown, M.L. Use of a post card query in mail surveys. Public Opinion Quarterly, 1965, 29, 635-637.

49. Brownlee, K.A. A note on the effects of nonresponse on surveys. American Statistical Association Journal, 1947, 42, 29-32.

50. Bruvold, W.H. Category and successive intervals scales for rating statements and stimulus objects. Journal of Experimental Psychology, 1969, 81, 230-235.

51. Bruvold, W.H. Reconciliation of apparent nonequivalence among alternative rating methods. Journal of Applied Psychology, 1977, 62, 111-115.

52. Cahalan, D. Effectiveness of a mail questionnaire technique in the army. Public Opinion Quarterly, 1951, 15, 575-578.

53. Campbell, D.T. The indirect assessment of social attitudes. Psychological Bulletin, 1950, 47, 15-38.

54. Campbell, D.T. From descriptions to experimentation: Interpreting trends as quasi experiments. In C.W. Harris (Ed.) Problems in measuring change. Madison: University of Wisconsin Press, 1963.

55. Campbell, D.T. & Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

56. Campbell, D.T. & Stanley, J.C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1966. Reprinted from N.C. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963.

57. Cannel, C. & Fowler, F. Comparison of a self-enumerative procedure and a personal interview: A validity study. Public Opinion Quarterly, 1963, 27, 250-264.

58. Cantril, H. Problems and techniques: Experiments in the wording of questions. Public Opinion Quarterly, 1940, 4, 330-332.

59. Center for Educational Research. The development of a survey instrument codebook. Kalamazoo, Michigan: The Center for Educational Research, Western Michigan University, 1977.

60. Champion, D.J. & Sear, A.M. Questionnaire response rate: A methodological analysis. Social Forces, 1969, 47, 335-339.

61. Clausen, J.A. & Ford, R.N. Controlling bias in mail questionnaires. American Statistical Association Journal, 1947, 42, 497-511.

62. Clunies-Ross, C.W. Simplification to survey statistics. Commentary (The Journal of the Market Research Society), 1967, 9, 68-76.

63. Cochran, W.G. Sampling techniques, (2nd ed.) New York: Wiley, 1963.

64. Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, 20, 37-46.

65. Conway, F. Sampling: An introduction for social scientists. London: George Allen and Unwin, 1967.

66. Cook, S.W. & Selltiz, C. A multiple-indicator approach to attitude measurement. Psychological Bulletin, 1964, 62, 36-55.

67. Cook, T. D. The potential and limitations of secondary evaluations. In M.W. Apple; M.J. Subkoviak; & H.S. Lufler, (Eds.) Educational evaluation: Analysis and Responsibility. Berkeley: McCutchen, 1974.

68. Cook, T.D. & Campbell D.T. The design and conduct of quasi-experiments and true experiments in field settings. In M.D. Dunnette, (Ed.), Handbook of industrial and organizational psychology. Chicago: Rand McNally, 1976.

69. Cooley, W.W. & Lohnes, P.R. Multivariate data analysis. New York: Wiley, 1971.

70. Corey, S.M. Signed versus unsigned attitude questionnaires. Journal of Educational Psychology, 1937, 28, 144-148.

71. Costner, H.L. Criteria for measures of association. American Sociological Review, 1965, 30, 341-353.

72. Crespi, I. What kinds of attitude measures are predictive of behavior. Public Opinion Quarterly, 1971, 35, 327-334.

73. Crowley, F. Compensation of subjects for participation in research. School and Society, 1958, 87, 430-431.

74. Crutchfield, R.S. & Gordon, D.A. Variations in respondents' interpretations of an opinion-poll question. International Journal of Opinion and Attitude Research, 1947, 1, 1-12.

75. Cuber, J.F. & Gerberich, J.B. A note on consistency in questionnaire responses. American Sociological Review, 1946, 11(2), 13-15.

76. Daniel, W. Nonresponse in sociological surveys. Sociological Methods and Research, 1975, 3, 291-307.

77. Davis, J.A. Elementary survey analysis. Englewood Cliffs: Prentice Hall, 1971.

78. Davis, R.A. & Barrow, E.L. A critical study of the questionnaire in education. Education Administration and Supervision, 1935, 21, 137-144.

79. Dawes, R.M. Fundamentals of attitude measurement. New York: Wiley, 1972.

80. Deming, W.E. On errors in surveys. American Sociological Review, 1944, 9, 359-369.

81. Department of Health, Education, and Welfare. Protection of human subjects (Technical amendment). Federal Register, 1975, 40(50), part 2, 11854-11858.

82. Desing, M.F. Suggestions to the novice in the mechanics of research. The School Review, 1941, 49, 206-212.

83. Deutscher, I. Physicians' reactions to a mailed questionnaire: A study in "resistentialism." Public Opinion Quarterly, 1956 20, 599-604.

84. Deutscher, I. Words and deeds: Social science and social policy. Social Problems, 1966, 13, 235-254.

85. Dillman, D. Increasing mail questionnaire response in large samples of the general public. Public Opinion Quarterly, 1972, 36, 254-257.

86. Dillman, D. & Frey, J. Contribution of personalization to mail questionnaire response as an element of a previously tested method. Journal of Applied Psychology, 1974, 59, 297-301.

87. DiVesta, F.J. A developmental study of the semantic structures of children. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 249-259.

88. Dollard J. Under what conditions do opinions predict behavior? Public Opinion Quarterly, 1948-49, 12, 623-632.

89. Donald, M. Implications of nonresponse for the interpretation of mail questionnaire data. Public Opinion Quarterly, 1960, 24, 99-114.

90. Dotson, L.E. & Summers, G.F. Elaboration on Guttman scaling techniques. In G.F. Summers, (Ed.), Attitude measurement. Chicago: Rand McNally, 1970.

91. Draper, N.R. & Smith, H. Applied regression analysis. New York: Wiley, 1966.

92. Duncan, O. A socio-economic index for all occupations. In A. Riess, Jr., et al. (Eds.), Occupations and social status, New York: Free Press of Glencoe, 1961.

93. Durant, H. & Maas, I. Who doesn't answer? British Psychological Society Bulletin, 1956, 29, 33-34.

94. Eckland, B.K. Effects of prodding to increase mail-back returns. Journal of Applied Psychology, 1965, 49, 165-169.

95. Eckland, B.K. Retrieving mobile cases in longitudinal surveys. Public Opinion Quarterly, 1968, 32, 51-54.

96. Edgerton, H.A.; Britt, S.H.; & Norman, R.D. Objective differences among various types of respondents to mailed questionnaires. American Sociological Review, 1947, 12, 435-444.

97. Edwards, A.L. Techniques of attitude scale construction. New York: Appleton-Century-Crofts, 1957.

98. Einhorn, H.J. Alchemy in the Behavioral Sciences. Public Opinion Quarterly, 1972, 36, 368.

99. Ellis, A. The validity of personality questionnaires. Psychological Bulletin, 1946, 43, 385-444.

100. Ellis, A. Questionnaire versus interview methods in the study of human love relationships. American Sociological Review, 1947, 12, 541-553.

101. Ellis, A. Questionnaire versus interview methods in the study of human love relations. II. Uncategorized responses. American Sociological Review, 1948, 13, 61-65.

102. Ellis, R.; Endo, C.M.; & Armer, J.M. The use of potential non-respondents for studying nonresponse bias. Pacific Sociological Review, 1970, 13, 103-109.

103. Erdos, P.L. How to get higher returns from your mail surveys. Printers' Ink, 1957, 258, 30-31.

104. Erdos, P.L. Successful mail surveys: High returns and how to get them. Printers' Ink, 1957, 258, 56-60.

105. Erdos, P.L. Professional mail surveys. New York: McGraw-Hill, 1970.

106. Everitt, B. Cluster analysis. London: Heinemann Educahand Books, 1974.

107. Ferber, R. The problem of bias in mail returns: A solution. Public Opinion Quarterly, 1948, 12, 669-676.

108. Ferber, R. Which--mail questionnaires or personal interviews? Printers' Ink, 1948, 222(7), 44-47, 61, 64-65.

109. Ferriss, A.L. A note on stiulating response to questionnaires. American Sociological Review, 1951, 16, 247-249.

110. Finifter, B.M. The generation of confidence: Evaluating research findings by subsample relication. In H.L. Costner, (Ed.), Sociological methodology: 1972. San Francisco: Jossey-Bass, 1972.

111. Fischer, R.P. Signed versus unsigned personal questionnaires. Journal of Applied Psychology, 1956, 30, 220-225.

112. Fishbein, M. (Ed.) Readings in attitude theory and measurement. New York: John Wiley, 1967.

113. Fishbein, J. & Ajzen, I. Belief, attitude, intention, and behavior: An introduction to theory and research. Reading, Massachusetts: Addision-Wesley, 1975.

114. Fleiss, J.L. Measuring nominal scale agreement among many raters. Psychological Bulletin, 1971, 76, 378-382.

115. Ford, N. The advance letter in mail surveys. Journal of Marketing Research, 1967, 4, 202-204.

116. Ford, N.M. Consistency of responses in a mail survey. Journal of Advertising Research, 1969, 9(4), 31-33.

117. Ford, R.N. & Zeisel, H. Bias in mail surveys cannot be controlled by one mailing. Public Opinion Quarterly, 1949, 13, 494-501.

118. Frazen, R. & Lazarsfeld, P.F. Mail questionnaire as a research problem. Journal of Psychology, 1945, 20, 293-320.

119. Freed, M.N. In quest of better questionnaires. Personnel and Guidance Journal, 1964, 43, 187-188.

120. Freeman, L.C. & Turkoz, A. Invalidity of indirect and direct measures of attitude toward cheating. Journal of Personality, 1960, 28, 443-447.

121. Fuller, C. Effect of anonymity on return rate and response bias in a mail survey. Journal of Applied Psychology, 1974, 59, 292-296.

122. Galtung, J. Theory and methods of social research. New York: Columbia University Press, 1967.

123. Gannon, M.; Northern, J.; & Carroll, S. Characteristics of non-respondents among workers. Journal of Applied Psychology, 1971, 55, 586-588.

124. Gardner, P.L. Scales and statistics. Review of Educational Research, 1975, 45, 43-57.

125. Garson, G.D. Handbook of political science methods. Boston: Halbrook Press, 1971.

126. Gerberich, J.B. A study of the consistency of informant responses to questions in a questionnaire. Journal of Educational Psychology, 1947, 38, 299-306.

127. Gerberich, J.B. & Mason, J.M. Signed versus unsigned questionnaire. Journal of Educational Research, 1948, 42, 122-126.

128. Gibson, F. & Hawkins, B. Interviews versus questionnaires, American Behavioral Scientist, 1968, 12.

129. Glass, G.V. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 5, 3-8.

130. Glass, G.V. & Hakstian, A.R. Measures of association in comparative experiments: Their development and interpretation. American Educational Research Journal, 1969, 6, 403-414.

131. Glass, G.V. & Stanley, J.C. Statistical methods in education and psychology. Englewood Cliffs, New Jersey: Prentice-Hall 1970.

132. Glasser, G.J. & Metzger, G.D. Random-digit dialing as a method of telephone sampling. Journal of Marketing Research, 1972, 9, 59-64.

133. Glock, C.Y. (Ed.) Survey research in the social sciences. New York: Russell Sage Foundation, 1967.

134. Goldstein, H. & Kroll, B.H. Methods of increasing mail response. Journal of Marketing, 1957, 22, 55-57.

135. Goode, W.J. & Hatt, P.K. Methods in social research. New York: McGraw-Hill, 1962.

136. Goodman, L.A. A general model for the analysis of surveys. American Journal of Sociology, 1972, 77, 1035-1086.

137. Goodman, L.A. Kruskal, W.H. Measures of association for cross classifications. American Statistical Association Journal, 1945, 49, 732-764.

138. Goodman, L.A. Snowball sampling. Annals of Mathematical Statistics, 1961, 32, 148-170.

139. Gordon, L.V. Validities of the forced-choice and questionnaire methods of personality measurements, Journal of Applied Psychology, 1951, 35, 407-412.

140. Green, B.L. Attitude measurement. In G Lindzey, (Ed.), Handbook of social psychology (Vol. 1). Cambridge, Massachusetts: Addison-Wesley, 1954.

141. Greenburg, A.A. A method of coding questionnaires in marketing surveys. Journal of Marketing, 1949, 14, 456-458.

142. Greenburg, A. & Manfield, M. On the reliability of mailed questionnaires in product tests. Journal of Marketing, 1957, 21, 342-345.

143. Gullahorn, J.T. & Gullahorn, J.E. Increasing returns from non-respondents. Public Opinion Quarterly, 1959, 23, 119-121.

144. Gullahorn, J.E. & Gullahorn, J.T. An investigation of the effects of three factors on the response to mail questionnaires. Public Opinion Quarterly, 1963, 27, 294-296.

145. Guttman, L. The Cornell Technique for scale and intensity analysis. Educational and Psychological Measurement, 1947, 7, 247-280. Reprinted in G.F. Summers, (Ed.), Attitude measurement. Chicago: Rand McNally, 1970.

146. Guttman, L. A basis for scaling qualitative data. American Sociological Review, 1944, 9, 139-150. Reprinted in G.F. Summers, (Ed.), Attitude measurement. Chicago: Rand McNally, 1970.

147. Hamel, L. & Reif, H. Should attitude questionnaires be signed? Personnel Psychology, 1952, 5, 87-90.

148. Hancock, J.W. An experimental study of four methods of measuring unit costs of obtaining attitude toward the retail store. Journal of Applied Psychology, 1940, 24, 213-230.

149. Hansen, M.H. & Hurwitz, W.N. The problem of nonresponse in sample surveys. American Statistical Association Journal, 1946, 46, 517-529.

150. Hansen, M.H.; Hurwitz, W.N.; & Madow, W.G. Sample survey methods and theory. New York: Wiley, 1953.

151. Harman, H.H. Modern factor analysis. Chicago: University of Chicago Press, 1977.

152. Heaton, E.E., Jr. Increasing mail questionnaire returns with a preliminary letter. Journal of Advertising Research, 1965, 5(4), 36-39.

153. Heise, D.R. The semantic differential and attitude research. In G.F. Summers, (Ed.), Attitude measurement, Chicago: Rand McNally, 1970.

154. Hinricks, J.R. Effects of sampling, follow-up letters, and commitments to participation on mail attitude survey response. Journal of Applied Psychology, 1975, 60, 249-251.

155. Hinricks, J. & Gatewood, R. Differences in opinion-survey response patterns as a function of different methods of survey administration. Journal of Applied Psychology, 1967, 51, 497-502.

156. Hobson, A. The use of the correspondence method in original research American Statistical Association Journal, 1916, 15, 210-218.

157. Hochstim, J.R. & Demetrios, A.A. Personal follow-up in a mail survey: Its contribution and its cost. Public Opinion Quarterly, 1970, 34, 69-81.

158. Hodgkinson, H. L. & Edelstein, S. Questionnaires: In fact there is an error. Educational Researcher, 1972, 1, 9-10.

159. Holdaway, E. Different response categories and questionnaire response patterns. Journal of Experimental Education, 1971, 40(2), 57-60.

160. Hoyt, J.S., Jr. Do quantifying adjectives mean the same thing to all people? Minneapolis: University of Minnesota, Agricultural Extension Service, 1972.

161. Hubbard, A.W. Phrasing questions. Journal of Marketing, 1950, 15, 48-56.

162. Hubbard, F.W. Questionnaires. Review of Educational Research, 1939, 9, 502-507.

163. Huff, D. How to lie with statistics. New York: Norton, 1954.

164. Huffman, H. Improving the questionnaire as a tool of research. The National Business Education Quarterly, 1948, 17, 15-18, 55-61.

165. Husek, T.R. Acquiescence as a response set and as a personality characteristic. Educational and Psychological Measurement, 1961, 21, 295-307.

166. Hyman, H. Survey design and analysis. New York: Free Press, 1955.

167. Jahoda, M.; Deutsch, M.; & Cook, S.W. Data collection: The questionnaire and interview approach. In Research methods in social relations with special reference to prejudice. New York: Dryden Press, 1951.

168. Johnson, W.R. et al. Effects of alternative positioning of open-ended questions in multiple-choice questionnaires. Journal of Applied Psychology. 59, 1974, April, 219-221.

169. Jourard, S. & Lasakow, P. Some factors in self-disclosure. Journal of Abnormal and Social Behavior, 1966, 1, 15-37.

170. Kanuk, L. & Berenson, C. Mail surveys and response rates: A literature review. Journal of Marketing Research, 1975, 12, 440-453.

171. Kassarjian, H. & Nakaniski, M. A study of selected opinion measurement techniques. Journal of Marketing Research, 1967, 4, 148-153.

172. Katz, D. The interpretation of survey findings. Journal of Social Issues, 1946, 2, 33-44

173. Kawash, M.B. & Aleamont, L.M. Effect of personal signature on the initial rate of return of a mailed questionnaire. Journal of Applied Psychology, 1971, 55, 589-592.

174. Keane, J.G. Low cost, high return mail surveys. Journal of Advertising Research, 1963, 3(3), 28-30.

175. Kelly, R.L. The revolt against questionable questionnaires. Association of American Colleges Bulletin, 1931, 17, 377-390.

176. Kendall, M.G. Rank correlation methods. (4th ed.) London: Griffin, 1970.

177. Kephart, W.M. & Bressler, M. Increasing the responses to mail questionnaires: A research study. Public Opinion Quarterly, 1958, 22, 123-132.

178. Kerin, R.A. & Peterson, R.A. Personalization, response anonymity and response distortion in mail surveys. Journal of Applied Psychology, 1977, 62, 86-89.

179. Kim, J. Predictive measures of ordinal association. American Journal of Sociology, 1971, 76, 891-907.

180. Kish, L. Survey sampling. New York: Wiley, 1965.

181 Kivlin, J.E. Contributions to the study of mail-back bias. Rural Sociology, 1965, 30, 323-326.

182. Klein, S.; Maher, J.; & Dunnington, R. Differences between identified and anonymous subjects in responding to an industrial opinion survey. Journal of Applied Psychology, 1967, 51, 152-160.

183. Knudsen, D.D.; Pope, H.; & Irish, D.P. Response differences to questions on sexual standards: An interview-questionnaire comparison. Public Opinion Quarterly, 1967, 31, 290-297.

184. Komorita, S.S. Attitude content, intensity, and the neutral points on a Likert scale. Journal of Social Psychology, 1963, 61, 327-334.

185. Koos, L.V. The questionnaire in education. New York: Macmillan, 1928.

186. Krejcie, R. & Morgan, D. Determining sample size for research activities. Educational and Psychological Measurement, 1970, 30, 607-610.

187. Kyburg, H.E., Jr. & Smokler, H.E. Studies in subjective probability. New York: Wiley, 1964.

188. Lake, D.G.; Miles, M.B.; & Earle, R.B.,Jr. Measuring human behavior: Tools for the assessment of social functioning. New York: Teachers College Press, 1973.

189. LaPiere, R.T. Attitudes vs. actions. Social Forces, 1934, 13, 230-237.

190. Lansdowne, J. The mailed questionnaire in panel research: Some empirical observations. Social Forces, 1970, 49, 136-140.

191 Larson, R.F. & Catton, W.R., Jr. Can the mail-back bias contribute to a study's validity? American Sociological Review, 1959, 24. 243-245.

192. Lawson, F. Varying group responses to postal questionnaires. Public Opinion Quarterly, 1949, 13, 114-116.

193. Lazarsfeld, P.F. The use of mail questionnaires to ascertain the relative popularity of network stations in family listening. Journal of Applied Psychology, 1940, 24, 802-816.

194. Lehman, E.C., Jr. Tests of significance and partial returns to mail questionnaires. Rural Sociology, 1963, 28, 284-289.

195. Leik, R.K. & Gover, W.R. Integrated approach to measuring association. In H.L. Costner, (Ed.). Sociological Methodology: 1971. San Francisco: Jossey-Bass, 1971.

196. Leonard, W.M., II. Basic social statistics. St. Paul: West Publishing, 1976.

197. Leslie, L. Increasing response rates to long questionnaires. The Journal of Educational Research, 1970, 63, 347-350.

198. Levine, S. & Gordon, G. Maximizing returns on mail questionnaires. Public Opinion Quarterly, 1959, 22, 568-575.

199. Light, R.J. Measures of response agreement for qualitative data: Some generalizations and alternatives. Psychological Bulletin, 1971, 76, 365-377.

200. Light, R.J. & Margolin, B.H. An analysis of variance for categorical data. American Statistical Association Journal, 1971, 66, 534-544.

201. Likert, R. A technique for the measurement of attitudes. Archives of Psychology, 1932, 140, (Monograph) Reprinted in part in G.F. Summers, (Ed.), Attitude measurement. Chicago: Rand McNally, 1970.

202. Lindsay, E.E. Questionnaires and follow-up letters. The Pedagogical Seminary, 1921, 28, 303-307.

203. Linn, L.S. Verbal attitudes and overt behavior: A study of racial discrimination. Social Forces, 1965, 43, 353-364.

204. Linsky, A.S. A factorial experiment in inducing responses to a mail questionnaire. Sociology and Social Research, 1965, 49, 183-189.

205. Linsky, A. Stimulating response to mailed questionnaires: A review. Public Opinion Quarterly, 1975, 21, 384-388.

206. Lissitz, R.W. & Green, S.B. Effect of the number of scale points on reliability: A Monte Carlo approach. Journal of Applied Psychology, 1975, 60, 10-13.

207. Loether, H.J. & McTavish, D.G. Descriptive and inferential statistics. Boston: Allyn and Bacon, 1976.

208. Lord, F.M. On the statistical treatment of football numbers. American Psychologist, 1953, 8, 750-751.

209. Longworth, D.S. Use of a mail questionnaire. American Sociological Review, 1953, 18, 310-313.

210. Lorge, I. The fundamental nature of measurement. In E.F. Lindquist (Ed.), Educational Measurement. Washington, D.C.: American Council on Education, 1951.

211. Magii, F.N.; Fotion, N.G.; & Gold, D. A mail-questionnaire adjunct to the interview. Public Opinion Quarterly, 1962, 26, 111-114.

212. Manniche, E. & Hayes, D. Respondent anonymity and data matching. Public Opinion Quarterly, 1957, 21, 384-388.

213. Martin, D.J. & McConnell, J.P. Mail questionnaire response induction: The effect of four variables on the response of a random sample to a difficult questionnaire. Social Science Quarterly, 1970, 51, 409-414.

214. Mason, W.S.; Dressel, R.J.; & Bain, R.K. An experimental study of factors affecting response to a mail survey of beginning teachers. Public Opinion Quarterly, 1961, 25, 296-299.

215. Matell, M.S. & Jacoby, J. Is there an optimal number of Likert scale items: Study I: Reliability and validity. Educational and Psychological Measurement, 1971, 31, 657-674.

216. Mayer, C.S. & Pratt, R.W., Jr. A note on nonresponse in a mail survey. Public Opinion Quarterly, 1966, 30, 637-646.

217. McDonagh, E.C. & Rosenblum, L.A. A comparison of mailed questionnaires and subsequent structured interviews. Public Opinion Quarterly, 1965, 29, 131-136.

218. Metzner, C.A. An application of scaling to questionnaire construction. American Statistical Association Journal, 1950, 45, 112-118.

219. Metzner, H. & Mann, F. A limited comparison of two methods of data collection: The fixed alternative questionnaire and the open-ended interview. American Sociological Review, 1952, 17, 486-491.

220. Michael, J.A. & Weinberger, J.A. Federal restrictions on educational research: Protection for research participants. Educational Researcher, 1977, 6(1), 3-7.

221. Michael, J.A. & Weinberger, J.A. Federal restrictions on educational research: Privacy Protection Study Commission hearing. Educational Researcher, 1977, 6(4), 15-18.

222. Mihailov, S. The indirect questionnaire as a tool for sociological research. Mens en Maatschappy, 1966, 41, 455-459.

223. Miller, D.C. Handbook of research design and social measurement. (2nd ed.) New York: David McKay, 1970.

224. Mitchell, W., Jr. Factors affecting the rate of return on mailed questionnaires. American Statistical Association Journal, 1939, 34, 683-692.

225. Monroe, C. & Finkner, A.L. Handbook of area sampling. Philadelphia: Chilton, 1959.

226. Mooren, R. & Rothney, J. Sampling problems in follow-up research. Occupations, 1952, 30, 573-578.

227. Mooren, R. & Rothney, J. Personalizing the follow-up study. Personnel and Guidance Journal, 1956, 34, 409-412.

228. Moser, C.A. & Kalton, G. Survey methods in social investigation. (2nd Ed.) New York: Basic Books, 1972.

229. Muehl, D. (Ed.) A manual for coders. Ann Arbor: Survey Research Center, University of Michigan, 1961.

230. Mueller, J.; Schuessler, K.; & Costner, H. Statistical reasoning in sociology. (2nd ed.) Boston: Houghton Mifflin, 1970.

231. Nemanich, D. Is keypunch verification really necessary? Public Opinion Quarterly, 1972, 36, 260-262.

232. Nevin, J.R. & Ford, N.M. Effects of a deadline and a veiled threat on mail survey responses. Journal of Applied Psychology, 1976, 61, 116-118.

233. Newman, W. Differences between early and late respondents to a mailed survey. Journal of Advertising Research, 1962, 2(2), 37-39.

234. Neyman, J. On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection. Journal of the Royal Statistical Society, 1934, 97, 558-606.

235. Nie, N.H.; Hull, C.H.; Jenkins, J.G.; Steinbrenner, K.; & Bent, D.H. Statistical package for the social sciences. New York: McGraw-Hill, 1975.

236. Nixon, J.E. The mechanics of questionnaire construction. Journal of Educational Research, 1954, 47, 481-487.

237. Noelle-Neumann, E. Wanted--rules for wording structured questionnaires. Public Opinion Quarterly, 1970, 34, 191-201.

238. Norman, R.D. A review of some problems related to the mail questionnaire technique. Educational and Psychological Measurement, 1948, 7, 235-247.

239. Norton, J. The questionnaire. National Educational Association Research Bulletin, 1930, 8, 1.

240. O'Brien, T.C. & Shapiro, B.J. Statistical significance--what? In R.E. Kirk. Statistical issues. Monterey: Brooks/Cole, 1972.

241. Olson, M.R. An application of matrix sampling to the method of paired comparisons. Journal of Educational Measurement. 1978, 15, 49-52.

242. O'Neill, H.W. Response style influence in public opinion surveys. Public Opinion Quarterly, 1970, 31, 95-102.

243. Oppenheim, A.N. Questionnaire design and attitude measurement. New York: Basic Books, 1966.

244. Orr, D. & Neyman, C.A. Considerations, costs, and returns in a large-scale follow-up study. The Journal of Educational Research, 1965, 58, 373-378.

245. Osgood, C.E.; Suci, G.H.; & Tannenbaum, P.H. The measurement of meaning. Urbana: University of Illinois Press, 1957.

246. Owens, W.A. Background data. In M.D. Dunnette (Ed.). Handbook of industrial and organizational psychology. Chicago: Rand McNally, 1976.

247. Pace, R.C. Factors influencing questionnaire returns from former university students. Journal of Applied Psychology, 1939, 23, 388-397.

248. Parry, H.J. & Crosley, H.M. Validity of responses to survey questions. Public Opinion Quarterly, 1950, 14, 61-80.

249. Parsons, R. & Medford, T. The effect of advance notice in mail surveys of homogeneous groups. Public Opinion Quarterly, 1972, 36, 258-259.

250. Parten, M. Surveys, polls, and samples: Practical procedures. New York: Harper and Bros., 1950.

251. Paterson, D.G. & Tinker, M.A. How to make type readable. New York: Harper and Bros., 1940.

252. Payne, S.L. Case study in question complexity. Public Opinion Quarterly, 1950, 13, 653-658.

253. Payne, S.L. Thoughts about meaningless questions. Public Opinion Quarterly, 1950, 14, 687-696.

254. Payne S.L. The art of asking questions. Princeton, N.J.: Princeton University Press, 1951.

255. Pearlin, L. The appeals of anonymity. Public Opinion Quarterly, 1961, 25, 640-647.

256. Perz, D.C. & Andrews, F.M. Detecting causal priorities in panel study data. American Sociological Review, 1964, 29, 836-848.

257. Perry, P. Election survey procedures on the Gallup Poll. Public Opinion Quarterly, 1960, 24, 531-542.

258. Platt, J.R. Strong inference. Science, 1964, 146, 347-353.

259. Politz, A. & Simmons, W. An attempt to get the "not at homes" into the sample without callbacks. American Statistical Association Journal, 1949, 44, 9-31

260. Pomeroy, W.B. The reluctant respondent. Public Opinion Quarterly, 1963, 27, 287-293.

261. Potter, D.R.; Sharpe, K.M.; Hendee, J.C.; & Clark, R.N. Questionnaires for research: An annotated bibliography on design, construction, and use. Research paper PNW-140. Portland, Oregon: USDA Forest Service, 1972.

262. Pucel, D.; Nelson, H.; & Wheeler, D. Questionnaire follow-up returns as a function of incentives and responder characteristics. Vocational Guidance Quarterly, 1971, 19, 188-193.

263. Raj, D. The design of sample surveys. New York: McGraw-Hill, 1972.

264. Rand Corporation. A million random digits with 100,000 normal deviates. Glencoe, Ill.: Free Press, 1955.

265. Reid, S. Respondents and nonrespondents to mail questionnaires. Educational Research Bulletin, 1942, 21, 87-96.

266. Reuss, C.F. Differences between persons responding and not responding to a mailed questionnaire. American Sociological Review, 1943, 8, 433-438.

267. Richardson, M.W. Note on Travers' critical review of the forced-choice technique. Psychological Bulletin, 1951, 48, 435-537.

268. Robin, S. A procedure for securing returns to mail questionnaires. Sociology and Social Research, 1965, 50, 24-35.

269. Robins, L.N. The reluctant respondent. Public Opinion Quarterly, 1963, 27, 276-286.

270. Robinson, R.A. How to design a mail survey. Printers' Ink, 1952, 239(9), 27-30.

271. Robinson, R.A. How to boost returns from mail surveys. Printers' Ink, 1952, 258(6), 35-37.

272. Robinson, R.A. & Agisim, P. Making mail surveys more reliable. Journal of Marketing, 1951, 15, 48-56.

273. Robinson, J.P. & Shaver, P.R. Measures of social psychological attitudes. Survey Research Center, Institute for Social Research: University of Michigan, 1973.

274. Robinson, W.S. The statistical measurement of agreement. American Sociological Review, 1957, 22, 17-25.

275. Roeher, A.G. Effective techniques in increasing response to mailed questionnaires. Public Opinion Quarterly, 1963, 27, 299-302.

276. Rollins, M. The practical use of repeated questionnaire waves. Journal of Applied Psychology, 1940, 24, 770-772.

277. Rosen, N. Anonymity and attitude measurement. Public Opinion Quarterly, 1960, 24, 675-678.

278. Rosen, H. & Rosen, R.A.H. The validity of "undecided" answers in questionnaire responses. Journal of Applied Psychology, 1955, 39, 178-181.

279. Rosenau, J.N. Meticulousness as a factor in the response to mail questionnaires. Public Opinion Quarterly, 1964, 28, 312-314.

280. Rosenberg, M. The logic of survey analysis. New York: Basic Books, 1968.

281. Rosenthal, R. Experimenter effects in behavioral research. New York: Appleton, 1966.

282. Roslow, S. & Blankenship, A.B. Phrasing the question in consumer research. Journal of Applied Psychology, 1939, 23, 612-622.

283. Schuman, H. The random probe--A technique for evaluating the validity of closed questions. American Sociological Review, 1966, 31, 218-222.

284. Schwirian, K.P. & Blaine, H.R. Questionnaire return bias in the study of blue-collar workers. Public Opinion Quarterly, 1966, 30, 651-653.

285. Scott, C. Research on mail surveys. Journal of the Royal Statistical Society, 1961, 124, 143-205.

286. Scott, F.G. Mail questionnaires used in a study of older women. Sociology and Social Research, 1957, 41, 281-284.

287. Scott, W.A. Reliability of content analysis: The case of no final scale coding. Public Opinion Quarterly, 1955, 19, 321-325.

288. Selltiz, C.; Johoda, M.; Deutsch, M.; & Cook, S.W. Research methods in social relations. New York: Holt, Reinhart, & Winston, 1959.

289. Selvin, H.C. A critique of tests of significance in survey research. American Sociological Review, 1957, 22, 519-527.

290. Selvin, H.C. The logic of survey analysis. (Monograph 12), Survey Research Center, University of California, Berkley, 1965.

291. Selvin, H.C. & Hirschi, T. Principles of survey analysis. New York: Free Press, 1973.

292. Selvin, H.C. Stuart, A. Data dredging procedures in survey analysis. American Statistician, 1966, 20(3), 20-23.

293. Sessions, F.R.; Epley, R.; & Moe, E.O. The development, reliability and validity of an all purpose optical scanner questionnaire form. Public Opinion Quarterly, 1966, 30, 423-428.

294. Shafer, E.L. & Hamilton, J.F. A comparison of four survey techniques used in outdoor recreation research. USDA Forest Service Research Paper NE-86.

295. Sharma, S. & Singh, Y.P. Does the colour pull responses? Manas: A Journal of Scientific Psychology, 1967, 14(2), 77-79.

296. Sharp, H. The mail questionnaire as a supplement to the personal interview. American Sociological Review, 1955, 20, 718.

297. Shaw, M.E. & Wright J.M. Scales for the measurement of attitudes. New York: McGraw-Hill, 1967.

298. Sheth, J.N. & Roscoe, A.M. Impact of questionnaire length, follow-up methods, and geographical location on response rate to a mail survey. Journal of Applied Psychology, 1975, 60, 252-254.

299. Shoemaker, D.M. An application of item-examiner sampling to scaling attitudes. Journal of Educational Measurement, 1971, 8, 279-282.

300. Shuttleworth, F.K. A study of questionnaire technique. Journal of Educational Psychology, 1931, 22, 652-658.

301. Shuttleworth, F.K. Sampling errors involved in incomplete returns to mail questionnaires. Journal of Applied Psychology, 1941, 25, 588-591.

302. Siegel, S. Nonparametric methods for the behavioral sciences. New York: McGraw-Hill, 1956.

303. Sills, D.L. International encyclopedia of the social sciences. New York: Free Press, 1968.

304. Silvey, J. Deciphering data. New York: Longman, 1975.

305. Simon, R. Response to personal and form letters in mail surveys. Journal of Advertising Research, 1967, 7, 28-30.

306. Sjoberg, G. A questionnaire on questionnaires. Public Opinion Quarterly, 1954, 18, 423-427.

307. Sjoberg, G. Ethics, politics, and social research. Cambridge, Massachusetts: Schenkman, 1967.

308. Sjoberg, G. & Nett, R. A methodology for social research. New York: Harper & Row, 1968.

309. Skipper, J.K., Jr. & Ellison, M.D. Personal contact as a technique for increasing questionnaire returns from hospitalized patients after discharge. Journal of Health and Human Behavior, 1966, 7, 211-214.

310. Sletto, R.F. Pretesting of questionnaires. American Sociological Review, 1940, 5, 193-200.

311. Slocum, W.L.; Empey, L.T.; & Swanson, H.S. Increasing response to questionnaires and structured interviews. American Sociological Review, 1956, 21, 221-225.

312. Slonim, M. Sampling: A quick, reliable guide to practical statistics. New York: Simon & Schuster, 1960.

313. Snelling, W.R. The impact of personalized mail questionnaire. The Journal of Educational Research, 1969, 63, 126-129.

314. Snider, J. & Osgood, C. The semantic differential: A sourcebook. Chicago: Adline Press, 1968.

315. Speer, D.C. & Zold, A. An example of self-selection bias in follow-up research. Journal of Clinical Psychology, 1971, 27, 64-68.

316. Stafford, J. Influence of preliminary contact on mail returns. Journal of Marketing Research, 1966, 3, 410-411.

317. Stanton, F. Notes on the validity of mail questionnaire returns. Journal of Applied Psychology, 1939, 23, 95-104.

318. Stephan, F.F. & McCarthy, P. Sampling opinions. New York: Wiley, 1958.

319. Stouffer, S.A. Social research to test ideas. New York: Free Press, 1962.

320. Suchman, E.A. An analysis of "bias" in survey research. Public Opinion Quarterly, 1962, 26, 102-111.

321. Suchman, E.A. & Guttman, L. A solution to the problem of question "bias." Public Opinion Quarterly, 1947, 11, 445-455.

322. Suchman, L.A. & McCandles, B. Who answers questionnaires? Journal of Applied Psychology, 1940, 24, 758-769.

323. Sudman, S. Applied sampling. New York: Academic Press, 1976.

324. Sudman, S. On sampling of very rare human populations. Journal of the American Statistical Association, 1972, 67, 335-339.

325. Sudman, S. Reducing the cost of surveys. Chicago: Adline, 1967.

326. Sudman, S. On the accuracy of recording on consumer panels: I and II. Journal of Marketing Research, 1964, 1, 14-20, 69-83.

327. Sudman, S. The uses of telephones directories for survey sampling. Journal of Marketing Research, 1973, 10, 204-207.

328. Sudman, S. & Bradburn, N.M. Response effects in surveys: A review and synthesis. Chicago: Adline, 1974.

329. Summers, G.F. (Ed.), Attitude measurement. Chicago: Rand McNally, 1970.

330. Tatsuoka, M.M. Multivariate analysis: Techniques for educational and sychological research. New York: Wiley, 1971.

331. Theil, H. On the estimation of relationships involving qualitative variables. American Journal of Sociology, 1970, 76, 103-153.

332. Thurstone, L.L. Attitudes can be measured. American Journal of Sociology, 1928, 33, 529-554. Reprinted in G.F. Summers (Ed.), Attitude measurement. Chicago, Rand McNally, 1970.

333. Toops, H.A. Validating the questionnaire method. Journal of Personnel Research, 1923, 2, 153-169.

334. Toops, H.A. The returns from follow-up letters to questionnaires. Journal of Applied Psychology, 1926, 10, 92-101.

335. Toops, H.A. The factor of mechanical arrangement and typography in questionnaires. Journal of Applied Psychology, 1937, 21, 225-229.

336. Torgerson, W.S. Theory and methods of scaling. New York: Wiley, 1958.

337. Travers, R.M.W. A critical view of the validity and rationale of the forced-choice technique. Psychological Bulletin, 1951, 48, 62-70.

338. Tufte, E.R. (Ed.) The quantitative analysis of social problems. Reading, Mass.: Addison-Wesley, 1970.

339. Tukey, J.W. Analyzing data: Sanctification or detective work? American Psychologist, 1969, 24, 83-91.

340. Tull, D.S. & Albaum, G.S. Survey research: A decision approach. New York: Intext Educational Publisher, 1973.

341. U.S. Bureau of Labor Statistics. Handbook of methods for surveys and studies. Bulletin 1711. Washington, D.C.: U.S. Government Printing Office, 1971.

342. U.S. Bureau of the Census. Bureau of the Census catalog. Washington, D.C.: U.S. Government Printing Office.

343. U.S. Department of Commerce. Indexes to survey methodology literature. Washington, D.C.: U.S. Government Printing Office, 1974.

344. U.S. National Center for Health Statistics. Programs and collection procedures. Vital and Health Statistics, Series 1. Washington, D.C.: U.S. Government Printing Office, 1963-present.

345. U.S. National Center for Health Statistics. Replication: An approach to the analysis of data from complex surveys. Vital and Health Statistics, Series 2, No. 14. Washington, D.C.: U.S. Government Printing Office, 1966.

346. Vincent, C.E. Socioeconomic status and familiar variables in mail questionnaire responses. American Journal of Sociology, 1964, 69, 647-653.

347. Waisanen, F.B. A note on the response to a mailed questionnaire. Public Opinion Quarterly, 1954, 18, 210-212.

348. Wallace, D. A case for--and against--mail questionnaires. Public Opinion Quarterly, 1954, 18, 40-52.

349. Wallis, W.A. & Roberts, H.V. The nature of statistics. New York: Free Press, 1962.

350. Wang, C.K.A. Suggested criteria for writing attitude statements. Journal of Social Psychology, 1932, 3, 367-373.

351. Watson, R. Investigations by mail. Market Research, 1937, 7, 11-17.

352. Webb, E.J.; Campbell, D.T.; Schwartz, R.D & Sechrest, L. Unobtrusive measures: Nonreactive research in the social sciences. Chicago: Rand McNally, 1972.

353. Weinberger, J.A. & Michael, J.A. Federal restrictions on educational research: A status report on the Privacy Act. Educational Researcher, 1977, 6(2), 5-8.

354. Weinberger, J.A. & Michael, J.A. Federal restrictions on educational research: A status report on the Buckley Amendment and Freedom of Information Act. Educational Researcher, 1976, 5(11), 3-8.

355. Wells, W.D. & Dames, J. Hidden errors in survey data. Journal of Marketing, 1962, 26(4), 50-54.

356. Wembridge, E.R. & Means, E.R. Obscurities in voting upon measures due to double-negative. Journal of Applied Psychology, 1918, 2, 156-163.

357. Whitney, D.R. & Feldt, L.S. Analyzing questionnaire results: Multiple tests by hypotheses and multivariate hypotheses. Educational and Psychological Measurement. 1973, 33, 365-380.

358. Williams, A.F. & Wechsler, H. The mail survey: Methods to minimize bias owing to incomplete response. Sociology and Social Research, 1970, 54, 533-535.

359. Winck, R.J. & Campbell, D.T. Proof? No. Evidence? Yes. The significance of tests of significance . The American Sociologist, 1969, 4, 130-143.

360. Woodward, J.L. & DeLott, J. Field coding versus office coding. Public Opinion Quarterly, 1952, 16, 432-436.

361. Wotruba, T.R. Monetary inducement and mail questionnaire response. Journal of Marketing Research, 1966, 3, 398-400.

362. Yates, F. Sampling methods for censuses and surveys. New York: Hafner, 1960.

363. Young, P.V. & Schmid, C.F. Scientific social surveys and research. Englewood Cliffs, New Jersey: Prentice Hall, 1966.

364. Young, R.A.; Holland, I.I.; & Gilmore, A.R. Getting better returns from mail questionnaires. Journal of Forestry, 1970, 68, 723-724.

365. Zimmer, H. Validity of extrapolating nonresponse bias from mail questionnaire follow-ups. Journal of Applied Psychology, 1956, 40, 117-121.

INDEX TO BIBLIOGRAPHY

Addressing the questionnaire, 3*, 155, 118

Advantages and disadvantages of questionnaires, 25, 36, 100, 118, 120, 126, 156, 167, 239, 288

Anonymity/confidentiality, 10, 38, 40, 41, 42, 44, 53, 70, 101, 111, 121, 127, 147, 178, 182, 183, 212, 214, 255, 277

Applications of questionnaires, 25, 185, 193, 211, 296, 351

Attitude measurement, 2, 37, 50, 51, 53, 66, 72, 79, 97, 112, 113, 140, 171, 188, 241, 243, 273, 277, 297, 329, 326

- Guttman scales, 90, 145, 146
- Likert scales, 184, 201, 215
- Semantic differential, 87, 153, 245, 314
- Thurstone scales, 332

Bibliographies, 27, 162, 170, 203, 238, 248, 261, 285, 342, 343

Coding, 59, 82, 92, 141, 229, 287, 360

Computer programming, 231, 235, 304

Costs, 344, 325

Cover letter, 13, 60, 103, 148, 204, 209, 272, 364

Early vs. late responses, 102, 107, 117, 181, 194, 216, 233, 285, 322

Ethics, 4, 5, 81, 220, 221, 307, 353, 354

Follow-up

- General, 6, 85, 117, 226, 244, 259, 268, 298
- Mail, 27, 61, 94, 134, 144, 154, 177, 198, 202, 205, 272, 276, 286, 322, 334, 358
- Personal contact, 143, 157
- Postcard, 38, 44, 104, 174
- Telephone, 94, 197, 198, 209, 315

*The numbers represent entries in the bibliography not pages in the narrative.

History, 239, 250

Interpretation of results, 54, 159, 160, 172, 278

Interview vs. mail survey, 52, 57, 100, 101, 108, 118, 128, 148, 167, 183, 211, 217, 219, 288, 294, 306

Motivation to respond

General, 61, 85, 155, 197, 204, 205, 209, 213, 224, 232, 247, 262, 271, 275, 276, 279, 309, 311
Use of money, 31, 46, 73, 148, 177, 271, 300, 361

Nonresponse

Analysis of, 76, 102, 365
Biases due to, 11, 17, 20, 45, 49, 78, 80, 89, 96, 118, 149, 162, 170, 190, 216, 238, 265, 266, 284, 285, 301, 315, 317, 320, 328, 347, 348
Causes, 28, 60, 83, 93, 121, 123, 173, 181, 192, 224, 260, 269, 346
Reduction of, 13, 26, 44, 85, 86, 95, 96, 103, 104, 107, 112, 140, 152, 170, 174, 188, 191, 198, 205, 259, 279, 311, 364

Personalizing the survey, 9, 86, 173, 178, 213, 227, 305, 309, 311, 313

Prequestionnaire contact, 1, 30, 104, 115, 143, 152, 249, 262, 316, 347

Pretest, 272, 310

Questions

Content, 167
Interpretation by respondents, 39, 72, 84
Type, 15, 21, 29, 139, 267, 337
Wording, 18, 24, 29, 35, 39, 58, 74, 159, 160, 161, 237, 252, 253, 254, 282, 321, 350, 356

Questionnaire

Appearance, 22, 251, 295, 335
Design, 19, 48, 82, 83, 103, 119, 155, 164, 168, 175, 218, 224, 236, 242, 270, 271, 293, 335, 358, 364
Length, 29, 197, 214, 298, 310

Reliability, 23, 75, 116, 118, 126, 142, 158, 160, 206, 215, 272, 287, 293, 326

Representativeness, 47, 93

Sampling, 63, 65, 110, 137, 138, 150, 154, 180, 186, 225, 226, 234, 257, 263, 264, 299, 312, 318, 323, 324, 327, 362

Stated vs. actual behavior, 13, 80, 84, 88, 169, 189, 203, 352

Statistics

Advanced, 33, 69, 91, 106, 151, 256, 330
Appropriate use of, 8, 62, 82, 98, 136, 163, 208, 240, 289, 292, 304, 339, 357, 359
Measurement properties of, 7, 124, 210
Measures of 64, 71, 114, 130, 137, 176, 179, 195, 199, 200, 274, 331
Textbook, 34, 77, 131, 187, 196, 207, 230, 302, 338, 349

Study design, 16, 43, 55, 67, 68, 119, 129, 190, 222, 246, 256, 257, 258, 280, 290, 304, 341, 344, 345

Textbooks

Social research, 32, 34, 112, 122, 125, 135, 288, 303, 307, 308, 319, 363
Survey, 12, 14, 27, 36, 105, 133, 166, 167, 223, 228, 243, 250, 254, 280, 291, 340

Validity, 15, 24, 43, 47, 53, 55, 57, 70, 75, 80, 94, 88, 99, 110, 120, 128, 139, 162, 165, 169, 172, 178, 189, 191, 203, 215, 242, 248, 278, 281, 283, 317, 320, 328, 333, 352, 355, 365

Additional copies of papers in this series are available upon request for a fee of $3.00. Previous papers in this series are listed below. Please direct your inquiries to:

Mary Anne Bunda, Series Editor
The Evaluation Center
College of Education
Western Michigan University
Kalamazoo, MI 49008

1. Evaluation Bibliography edited by Mary Anne Bunda

2. Hints for Planning and Conducting a Survey and a Bibliography of Survey Methods by Randy E. Demaline and D. William Quinn